Education Tech Research Dev (2021) 69:809–841
https://doi.org/10.1007/s11423-020-09876-0

DEVELOPMENT ARTICLE

Simulation research and design: a dual-level framework for multi-project research programs

Maximilian C. Fink[1] · **Anika Radkowitsch**[2] · **Elisabeth Bauer**[2] · **Michael Sailer**[2,3] · **Jan Kiesewetter**[1,3] · **Ralf Schmidmaier**[1,3,4] · **Matthias Siebeck**[1,3] · **Frank Fischer**[2,3] · **Martin R. Fischer**[1,3]

Accepted: 6 November 2020 / Published online: 17 December 2020

Abstract

Collaborations between researchers and practitioners have recently become increasingly popular in education, and educational design research (EDR) may benefit greatly from investigating such partnerships. One important domain in which EDR on collaborations between researchers and practitioners can be applied is research on simulation-based learning. However, frameworks describing both research and design processes in research programs on simulation-based learning are currently lacking. The framework proposed in this paper addresses this research gap. It is derived from theory and delineates levels, phases, activities, roles, and products of research programs to develop simulations as complex scientific artifacts for research purposes. This dual-level framework applies to research programs with a research committee and multiple subordinate research projects. The proposed framework is illustrated by examples from the actual research and design process of an interdisciplinary research program investigating the facilitation of diagnostic competences through instructional support in simulations. On a theoretical level, the framework contributes primarily to the literature of EDR by offering a unique dual-level perspective. Moreover, on a practical level, the framework may help by providing recommendations to guide the research and design process in research programs.

Keywords Educational design research · Simulation-based learning · Research–practice partnerships

---

✉ Maximilian C. Fink
maximilian.fink@med.uni-muenchen.de

[1] Institute for Medical Education, University Hospital, LMU Munich, Pettenkoferstraße 8a, 80336 Munich, Germany

[2] Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

[3] Munich Center of the Learning Sciences, Ludwig-Maximilians-Universität München, Munich, Germany

[4] Department of Medicine IV, University Hospital, LMU Munich, Munich, Germany

Introduction

Research–practice partnerships have recently become more widespread in education (Coburn and Penuel 2016). Such partnerships frequently strive to close the gap between a steadily growing body of research and a lack of improvement in educational practice. Specifically, research–practice partnerships address this gap by developing and disseminating solutions to highly-relevant educational problems and creating practical knowledge applicable to real-life contexts (Butler 2008; Ormel et al. 2012). Educational design research (EDR) focusses on the creation and evaluation of specific interventions as well as the scientific exploration of the design and production process itself (Design-Based Research Collective 2003). EDR and similar approaches may thus benefit greatly from investigating close collaborations between researchers and practitioners and could potentially optimize these collaborations.

One domain in which EDR on collaborations between researchers and practitioners can be applied is research on simulation-based learning. In this research area, intriguing research questions on learning and assessment remain unresolved (Heitzmann et al. 2017), and research programs can develop valuable educational products. Research programs on simulation-based learning possess three crucial features that could also characterize other interesting research–practice partnerships. First, research is carried out in (interdisciplinary) teams. Second, research programs collaborate on a management level as well as on the level of individual projects. Third, research and design coincide within at least a single phase of these programs, due to the custom development of complex products (such as simulation components) over a long period.

A framework for research and design in research programs on simulation-based learning is lacking. Such a framework is an essential complement to the scarce literature on the interplay between research and design in the field of EDR, which has not yet sufficiently delineated the phases, activities, and products of research programs that create complex scientific artifacts (McKenney and Reeves 2012). In particular, the current literature does not include a dual-level, multi-project framework for collaborative research and design (i.e., a model that delineates the role of a research committee and various subordinate research projects). Such a framework could support researchers and designers in creating effective research materials and simulations. Moreover, it could aid team leaders in systematically monitoring and guiding personnel and their collaboration.

In this article, we will first outline the relevant theoretical background and then propose a dual-level framework for multi-project research programs on simulation-based learning. This framework outlines phases, activities, roles, and products of simulation research and design. Notably, it applies to research programs with a research committee and multiple subordinate research projects focusing on the assessment and facilitation of knowledge and competences. We offer recommendations for each level and phase of the framework and illustrate the proposed framework with examples from the actual research and design process of the research program COSIMA. This research program investigates the use of simulations to facilitate diagnostic competences in the domains of medical education and teacher education. In the discussion, we will summarize our findings, address the implications for theory and practice, discuss the limitations of the framework, and provide directions for future research.

Theoretical background

The theoretical background of our framework is grouped into two parts. In the first part, we review the literature on research and design in educational research programs. Due to its focus on all types of educational research programs, this section is relatively broad in scope and more general. In the second part, we summarize simulation research and design in medical education and teacher education. Thus, the scope of the second part is more narrow due to its domain-specific focus. Together, the two parts complement one another, summarizing both the general and domain-specific theoretical background of our framework.

Research and design in educational research programs

We first outline the characteristics and stages of research programs, before elucidating relevant frameworks from educational design research. Finally, we describe important roles in educational research and design.

Characteristics and stages of research programs

According to the team science literature, research programs can be characterized as multi-team systems in which research projects collaborate and work autonomously to achieve common objectives (Shuffler et al. 2015). Similar to units in other contexts, these multi-team systems can be organized in a rather flat or strongly hierarchical structure and can include units of different sizes stemming from the same or different institutions (DeCostanza et al. 2014). Research programs often have a relatively strong hierarchical structure and consist of a research committee and subordinate research projects (Bozeman and Boardman 2014). While the research committee performs leadership tasks such as creating a common vision (Gray 2008) and answering overarching research questions, the subordinate research projects investigate more specific research questions. Next, we present the stages that typically take place within research projects before discussing this point specifically with respect to research programs.

Traditional models delineated research, regardless of the specific context, as a linear process comprised of clearly separable stages (e.g., Finley and Pocoví 2000). Such stages include a literature review, formulation of hypotheses, conducting empirical investigations, analyzing data and communicating findings. Research process models like this can be used to describe individual *research projects* and have been applied in multiple disciplines to investigations with various methodologies. Unfortunately, these models barely resemble scientists' actual, rather complex, research process. Contemporary models depict the research process more accurately as the adaptive and iterative application of activities (e.g., evidence generation, drawing conclusions, communicating and scrutinizing findings) to the current status of an operation (Fischer et al. 2014; Reiff et al. 2002). These activities are relatively specific and can occur in various research contexts, including individual *research projects* as well as *research programs*. In a model of collaborative research, Sonnenwald (2007) described that *research programs* typically traverse through the stages of foundation, formulation, sustainment, and conclusion. The foundation stage designates conditions before and during the start of a collaboration and includes but is not limited to resource availability and building professional networks. In the formulation stage, collaborators devise a research program and apply for a grant. During the sustainment stage, the research program that has received funding and started its activities is evaluated continuously and kept on track. In the conclusion stage, outcomes of the research program are published, and new frameworks and proposals are created.

In summary, it can be said that research programs often consist of a research committee and multiple individual, subordinate research projects. Likewise, the research process in this specific context can be described quite well by the stages mentioned by Sonnenwald (2007). The new framework proposed in this article will thus characterize research programs as dual-level, multi-team systems and contain phases based on the stages delineated by Sonnenwald (2007).

Educational design research

The term EDR subsumes multiple approaches, such as design-based research, that simultaneously pursue the objectives of creating scientific knowledge and designing educational artifacts (McKenney and Reeves 2012). However, in the context of EDR, there are only a few models and frameworks that adequately represent research and design in research programs. Next, we will provide a brief overview of these models and frameworks.

Bannan-Ritland (2003) proposed one of the first EDR models to address research and design on a *research program* level, the integrative learning design framework (ILDF). The ILDF comprises the phases of exploration, enactment, and evaluation. This framework encompasses a rapid prototyping approach and aims for the adoption of the designed product. Middleton et al. (2008) presented an innovative EDR model for *research projects* with the “compleat” design experiment. According to the “compleat” design experiment, an empirical investigation (e.g., a randomized controlled trial) is designed and developed as part of an educational research process. A feasibility study, rapid prototyping, and pilot study are significant steps within the design and validation of the empirical investigation. The osmotic model (Ejersbo et al. 2008) combines a research process model for *research projects* with a product design model. This model depicts the interaction between research and design as ideal if both processes take place simultaneously and provide necessary input for one another. According to another EDR model by Akkerman et al. (2013), the three epistemic practices of research, design, and educational change take place concurrently in single *research project* EDR. The model highlights that each of these epistemic practices involves different subjects (e.g., researcher, designer, and change agent), tools (e.g., specific underlying models) and outcomes (e.g., a publication, a developed application or a developed application integrated into a curriculum). Synthesizing different EDR models, the generic model for EDR (McKenney and Reeves 2012) delineates the prototypical phases in which scientific knowledge and an educational intervention are produced and refined. The model depicts research and design as of equal status and posits that EDR processes in single *research projects* follow the iterative phases of (1) analysis and exploration, (2) design and construction, and (3) evaluation and reflection. During these phases, an educational artifact and scientific knowledge are developed and translated to practice.

In short, the described models and frameworks highlight essential phases and products of EDR. The new framework proposed in this paper will describe an EDR process in a research program. EDR processes in research programs have already been described by the ILDF (Bannan-Ritland 2003). However, our framework goes beyond this and other EDR frameworks by offering a separate description for the research committee and research project level as well as a context-specific focus on simulation research and design. Moreover, our framework will embed elements of the generic model for EDR (McKenney and

Reeves 2012), which emphasizes an equal status of research and design. An equal status of research and design is crucial in one phase of our framework when research materials and simulation components are developed and evaluated simultaneously. We will, thus, incorporate the design and construction, and evaluation and reflection phases of the generic model for EDR in this phase of our framework. In the next section, we will discuss major roles that play a part in educational research and design.

Roles in educational research and design

Throughout the educational research and design process, team members can fulfill multiple roles, including consultant, researcher, and designer (McKenney and Brand-Gruwel 2018). In the consultant role, team members mainly provide training, access to professional networks, and (scientific) advice to team members. In the researcher role, team members conduct (empirical) studies or carry out literature research to gain knowledge or develop theory. In the designer role, team members specify and develop artifacts. The developed artifacts are evaluated and improved until they meet the specified requirements. Apart from these roles, content experts fulfill an important role in educational research and design processes (Lee 1994) because they choose, create and evaluate the content of educational artifacts based on their extensive professional knowledge. The new framework presented in this paper will include all of these described roles, not least because these roles seem to be highly applicable for the context of research on simulation-based learning.

Simulation research and design in medical education and teacher education

An integrative perspective on simulation research and design in medical education and teacher education may enable generalization to other contexts and is warranted for three key reasons. First, simulations act as important "approximations of practice" (Grossman et al. 2009, p. 2058) within college courses in both domains and can be used for similar training and assessment purposes (Kaufman and Ireland 2016; Ryall et al. 2016). Second, some joint conceptualizations of knowledge and competences already exist. This is particularly the case for diagnostic competences (Förtsch et al. 2018; Heitzmann et al. 2019). Third, we can assume that the instructional design of simulations in both domains is rather similar. This should be the case because similar design processes emerge for creating interactive simulations in medicine and teacher education (Dotger et al. 2010), and comparable design features (e.g., feedback and clear educational aims) must be considered in both domains (Badiee and Kaufman 2015). Furthermore, we believe that the instructional design process should be particularly similar in both domains if we focus on simulations that assess and train knowledge and competences and exclude simulations that emphasize assessing and training other types of skills, such as motor skills. In the following section, we outline important features and terms in simulation-based learning.

Features and terms in simulation-based learning

Simulations can be defined as authentic models of professional situations that can be manipulated by participants (Jones et al. 2015; Kaufman and Ireland 2016). In contrast, *simulators* refer to (technical) devices and environments deployed to conduct simulations (Khan et al. 2011). This distinction of terms is meaningful for the design process because it demonstrates that simulations and simulators are distinct artifacts that have to be designed separately. Above all, simulations can be characterized based on their modality (Ziv et al. 2000). *Live simulations* involve a professional standardized person such as a standardized patient or student who has received training and prepared diligently for the case (Barrows and Abrahamson 1964). In *role-play simulations*, participants are usually only prepared through a short training phase and receive a supporting script before they interact with each other in different roles (e.g., physician and patient; Gartmeier et al. 2015; Simpson 1985). *Digital simulations* are conducted using a computer and a virtual person or environment to interact with (de Jong 1991). Alternatively, simulations can be characterized by their purpose, either assessment, facilitation, or research (Crawford 1966). As targeted research and design processes should consider the critical features of modality and purpose, the new framework proposed in this article will provide specific recommendations for these two features. Moreover, our framework will address the assessment and conceptualization of knowledge and competences in medical education and teacher education.

Assessment and conceptualization of knowledge and competences

Knowledge has long been assessed in medical education and teacher education through traditional paper–pencil tests using constructed response (e.g., short answer questions) or closed response formats (e.g., multiple-choice questions; Kastner and Stangla 2011). Nowadays, however, knowledge is increasingly assessed using such formats as part of simulation-based learning. Knowledge tests are therefore frequently integrated into testing environments before, during, or after simulation-based assessment. The most common knowledge classification in teacher education (Shulman 1987) categorizes knowledge based on content, while a popular knowledge classification in medical education also makes categorizations based on structure (Paris et al. 1983). Based on knowledge classifications from both fields, Förtsch et al. (2018) proposed an interdisciplinary framework for medical education and teacher education in which knowledge can be differentiated according to *types of knowledge* (i.e., the structure of knowledge) as well as *content-related facets* of knowledge. In contrast to knowledge, competences can be assessed rather well directly within simulations by observing participants' demonstrated performance (Blömeke et al. 2015; Miller 1990). Assessment can also take the form of multiple situation evaluations in which participants move from one evaluation station to the next (Harden et al. 1975). The result and process of participants' demonstrated performance are often evaluated in such assessments by raters applying rating scales (Rothman et al. 1996) and evaluated in comparison to solutions by one or even multiple experts (Charlin et al. 2010). The new framework presented in this article will offer specific recommendations for assessing knowledge and competences based on the presented literature from both domains. We discuss different frameworks for research with simulations below.

Research with simulations

To our knowledge, no comprehensive framework for multi-project research programs on simulation-based learning is available. However, an existing framework for multi-center simulation research (Cheng et al. 2017) provides some insights that are also applicable to research programs. According to this framework, multi-center simulation research is conducted in four separate but overlapping phases, in which the project is first planned and developed, before a study is conducted and findings are communicated. More literature is available on specific topics within simulation research, such as creating valid and reliable simulation scenarios and instruments, standardizing studies, recording, evaluating and reporting data and adherence to research ethics (Cheng et al. 2014a, 2016; Lamé and Dixon-Woods 2018). In addition, simulations have been used in a large number of studies to conduct basic research (Cook et al. 2013; Kaufman and Ireland 2016) as well as EDR (e.g., de Coninck et al. 2019; Hirumi et al. 2016a; Koivisto et al. 2018). These studies demonstrate that materials developed for research purposes include *instruments that assess knowledge and competences*, an *educational intervention*, and an *experimental procedure* (see Table 1 for our *own* definitions of these terms). The phases of research with simulations described by Cheng et al. (2017) and the classification of research materials in Table 1 will be included in the new framework presented in this article.

Design of simulations

Currently, only limited literature is available regarding the design of simulations. Cheng et al. (2014a) provided an overview of simulation research and design in pediatrics, including design recommendations for selecting an appropriate simulator, designing *simulation scenarios* (see Table 2 for the definition), and establishing sufficient authenticity. In the framework for multi-center simulation research mentioned above, Cheng et al. (2017) point out that researchers should scrutinize log data and evaluate the simulation scenarios created through pilot studies and with institutional review committees. Apart from these recommendations, there are only frameworks and recommendations for the design of specific types of simulations, such as live simulations (Khan et al. 2013a, b; Sturpe and Schaivone 2014) and digital simulations (Posel et al. 2015; Zary et al. 2006), as well as specific design activities such as scenario design (Benishek et al. 2015). Regarding the simulation design process in research programs, a number of instructional design studies from teacher education (e.g., Christensen et al. 2011; Ferry and Kervin 2007) and medical education (e.g., Hirumi et al. 2016b; Jensen et al. 2015) illustrate that various simulation components have been developed. These simulation components typically include *learning and testing environments*, *simulation scenarios*, *case vignettes*, as well as *briefings* and *debriefings* (see Table 2 for definitions of these terms). The classification of simulation components in Table 2 will be incorporated into the new framework proposed in the next section.

Table 1 Overview of research materials

Term	Definition
Instruments that assess knowledge and competences	Instruments that assess knowledge focus on knowledge as the prerequisite of performance or learning outcome. Instruments that assess competences operationalize this construct as observable results and processes in simulations. Such instruments can include (computer-based) tests, coding schemes, and rating scales
Educational intervention	The treatment condition examined in an empirical investigation (e.g., different kinds of instructional support, such as prompts, that may foster knowledge and competences)
Experimental procedure	The experimental procedure consists of the measurement procedure and an experimenter's guide
Measurement procedure	The measurement procedure is a report of all used instruments and the sequence of measurements during an empirical investigation
Experimenter's guide	The experimenter's guide is a specification for experimental standardization and may include (verbal) instructions

Table 2 Overview of simulation components

Term	Definition (and source)
Learning environment	Learning environments pursue the goal of facilitating knowledge and competence acquisition and include instructional support. An example of such an environment is an e-learning platform involving elaborate feedback
Testing environment	Testing environments pursue the goal of assessing knowledge, competences, and other variables and do not contain instructional support. An example of such an environment is an e-learning platform used in exams
Simulation scenario	The professional situation in which an agent (such as a learner acting as teacher or physician) can demonstrate competences, its characteristics, and development over time (Huffman et al. 2016)
Case vignette	A case that depicts a person (such as a patient or student) or multiple persons with specific characteristics and contains a particular task in a simulation scenario
Briefings	Briefings typically include a familiarization as well as a fiction contract. Briefings support learners in finding their way into the simulation and create a sense of safety (Rudolph et al. 2014)
Debriefings	Debriefings are a learning opportunity in which learners compare their solution to the simulation's designated solution. Debriefings often include feedback or group reflection (Fanning and Gaba 2007). Moreover, debriefings can be used to discharge actors in live simulations gently from the situation

A dual-level framework for simulation research and design in multi-project research programs

Overview of the framework

In line with the notion of hierarchical, multi-team systems (Bozeman and Boardman 2014; DeCostanza et al. 2014; Shuffler et al. 2015), we propose a framework that describes a research program as a dual-level, multi-project system consisting of one research committee and two or more research projects. Based on the stages of research programs (Sonnenwald 2007), and the phases of the multi-center model for conducting simulation-based research programs (Cheng et al. 2017), we posit that the research committee and research projects pass through the following program/project phases: 1. Creating a foundation for a research program or project; 2. Constructing and adapting simulations and research materials; 3. Conducting empirical investigations; and 4. Analyzing data and communicating results. Decisions in the first program/project phase affect all subsequent process phases. During the second project phase, the research materials and simulation components listed in Tables 1 and 2 are constructed and adapted, following the *design and construction* as well as *evaluation and reflection* phases of the generic model for EDR (McKenney and Reeves 2012). Products developed within the research program include scientific knowledge as well as practical applications and tools. Even though earlier program/project phases create research materials and simulation components as well as products for subsequent program/project phases, they are not necessarily executed as separate phases in a strictly linear format. Rather, program/project phases can overlap with each other and go through multiple iterations, drawing on previously developed research materials, simulation components, and products. For instance, one program/project phase can start before another has been completed. Members of the research committee and the research projects take on the roles of consultant, researcher, designer (McKenney and Brand-Gruwel 2018), and content expert (Lee 1994) throughout the program/project phases. While team members on both levels can fill multiple roles at the same time, certain roles can also be stressed for specific members in particular program/project phases. The described framework is illustrated for one research committee and multiple fictional research projects in Fig. 1.

Below, the four separate program/project phases will be explained in terms of the two levels of the framework. Relevant research materials and simulation components were defined in Tables 1 and 2. Recommendations for and an overview of the activities in the different phases will be presented for the research committee level in Table 3 and for the research project level in Table 4. Concurrently to our explanation of the program/project phases, we will illustrate the proposed framework by reporting on the research process that took place in the research program COSIMA. This research program investigates facilitating diagnostic competences with scaffolding in simulation-based learning environments in higher education and consists of a research committee and eight research projects. Seven empirical research projects gather data through experiments in the domains of medical education and teacher education; one meta-analysis project synthesizes data. A research committee leads and coordinates activities between the research projects. As an example research project, we will report on research project COSIMA 4 that investigates facilitating interactive diagnostic competences in simulated clinical history-taking.

Phase 1: creating a foundation for a research program or research project

Research committee level

At the beginning of this phase, suitable collaborators with common interests are identified, form a team, and apply for a grant posing an overarching research question for the prospective research program (Cheng et al. 2017; Sonnenwald 2007). The research committee's proposal may include a preliminary version of a conceptual framework, a common methodology, and a strategy to synthesize data from different research projects. After program start, the research committee states common objectives for all projects, coordinates activities between projects, assigns tasks, provides background knowledge on the topic and establishes a common terminology (Beck et al. 2017; Bennett and Gadlin 2012; Hall et al. 2012).

Research project level

In this phase, each research project creates a research proposal. In order to do so, the project first poses a research question and develops a research design. Then, the project determines a professional situation to simulate with high practical relevance, operationalizes relevant knowledge and competences, and chooses a suitable simulator and instructional support (see details and recommendations on these activities in Table 4). Similar to the analysis and exploration phase of the generic model for EDR (McKenney and Reeves 2012), researchers concurrently gain a theoretical understanding of the problem and determine possible solutions. In the current context, expert interviews and literature reviews are carried out to gain first insights into relevant variables and effective educational interventions for the selected simulated scenario and topic. Researchers, designers, and content experts also evaluate similar studies and commercial products to decide, with input from

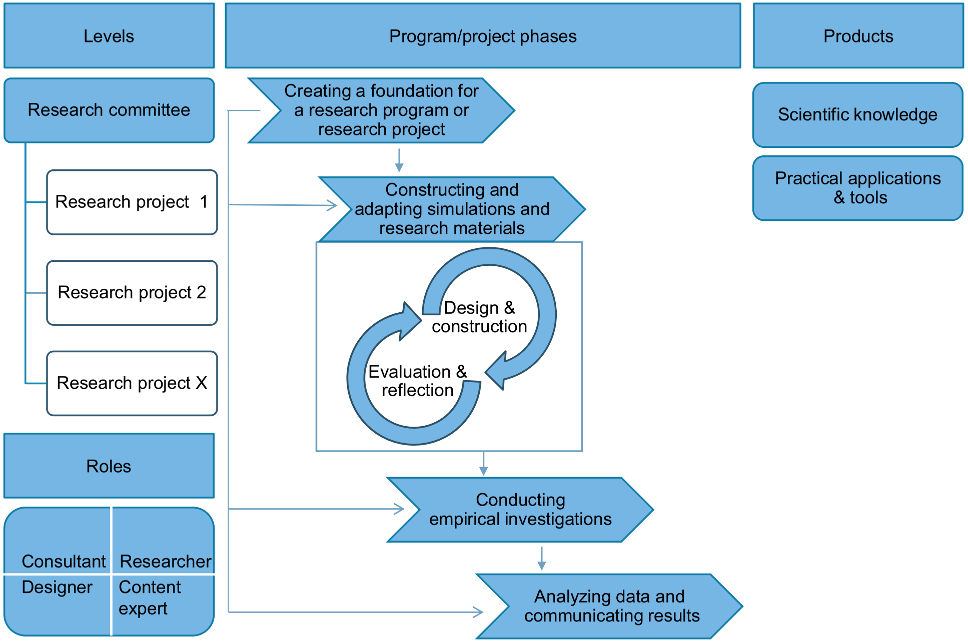


Fig. 1 A dual-level framework for simulation research and design in multi-project research programs

Table 3 Checklist for conducting research and design on the research committee level

Phase	Activities	Recommendations and important literature
Creating a foundation for a research program	Identify collaborators	Search for collaborators with common research interests and acknowledged expertise. Also, collaborators should possess enough time and be able to work together as a team
	Develop an overarching research question	Choose an important, rather broad research question that can only be answered conclusively by a research program consisting of multiple projects
	Develop a research proposal	This proposal should connect all projects. It could include a preliminary framework, common methodology, overview of joint research questions, and elaboration of supervision and publication strategies. Choose similar designs and research questions in several research projects to allow for generalization across contexts and domains. See general recommendations for grant writing (Gitlin and Lyons 2013) and specific recommendations for research programs (Cheng et al. 2017)
Constructing and adapting simulations and research materials	Create requirements for research and design	Set specific requirements for the learning and testing environments, scenarios, case vignettes, and (de)briefings (see Table 2). Provide a battery of common instruments and specify the intervention and measurement procedure (see Table 1)
	Monitor the fulfillment of requirements	Monitor what projects have fulfilled which requirements. Screen for problems that occur in multiple projects and take counter-measures
	Provide guidance to researchers and designers	Offer regular meetings in which members of research projects can clarify questions and get support in fulfilling requirement
	Convey a common knowledge base	Convey knowledge on the developed research materials and simulation components and provide examples from previous research or the literature. Convey basic knowledge on the instructional design process in general (Gagné 2007) and on simulation design specifically (Cheng et al. 2014a, 2017)
Conducting empirical investigations	Specify final targets for the empirical investigations	Determine which final targets the empirical investigations have to fulfill to answer the research committee's research questions (e.g., specify how many participants have to complete a certain instrument for an overarching data analysis of multiple projects)

Table 3 (continued)

Phase	Activities	Recommendations and important literature
	Monitor the standardization across research projects	Make sure that comparable aspects of the experimental procedure are carried out similar in research projects (e.g., randomization processes should be standardized, Cheng et al. 2017)
	Monitor the empirical investigations and resolve co-occurring problems	Track which projects encounter problems and whether these problems co-occur in other projects. Conduct site visits and check log data
	Provide guidance	Offer counseling to research projects regarding individual problems
	Convey a common knowledge base	Convey knowledge on conducting research with simulations, recruitment, and recording data
Analyzing data and communicating results	Provide specifications on reporting data and gather data of the research projects	Provide reporting guidelines for simulation research (Cheng et al. 2016). Ask each research project to provide full and standardized data documentation. Gather and merge all data systematically
	Analyze gathered data and convey knowledge	Use special methods to analyze large data sets from multiple projects. Such methods can include, for example, meta-analytic methods and structural equation modelling. Organize methods workshops for the projects (e.g., on mining learning processes in simulations)
	Create scientific publications independent from project data	Write theoretical articles, methodological articles, literature reviews, or meta-analyses. Such publications can be used as theoretical background in research projects if they are available when these projects reach the phase of analyzing data and communicating results
	Create scientific publications that depend on project data	Create publications with a common topic based on several research projects or that go further than the research projects in terms of a particular methodological aspect
	Increase visibility of findings	Develop a social media strategy and a strategy to disseminate findings in popular media formats (e.g., on Twitter or ResearchGate)
	Generate first ideas for an extension grant	Start developing first ideas for an extension grant with sufficient time. Analyze what innovative research questions still have to be answered and what materials can be reused

Table 4 Checklist for conducting research and design on the research project level

Phase	Activities and subactivities	Recommendations and important literature
Creating a foundation for a research project	Pose a research question	Choose a focused research question that can be answered by a single research project
	Develop the design	Determine dependent and independent variables; decide on a longitudinal or cross-sectional design. Keep in mind that one should have an appropriate number of simulated cases in the pretest, intervention, and posttest to reach sufficient reliability. Calculate the required sample size and make sure to have sufficient statistical power
	Select a professional situation to simulate	Select a practically relevant professional situation. Frequent situations, situations with serious consequences, and situations that are difficult or expensive to train are examples
	Operationalize relevant knowledge and competences	Determine what knowledge and which competences are relevant in the professional situation. Select the most suitable conceptual frameworks and assessment instruments
	Choose a suitable simulator	Choose a simulator that provides the necessary key features such as sufficient authenticity and interactivity and detailed log data. Consider the necessary budget and maintainability
	Select instructional support	Select a type of instructional support (e.g., sample solutions) that does not intrude on the intervention. Deliberately select the level of instructional support in all conditions
	Gain a theoretical understanding of the problem and determine possible solutions	Interview experts, conduct site visits and review the literature to improve your understanding of the simulated situation, intervention, and simulators. Determine possible solutions based on your findings from these sources
Constructing and adapting simulations and research materials		
Design and construction	Decide on a learning and testing environment	Evaluate what learning and testing environments are available and within budget. Consider simulator integration, available test formats, quality of log data, user administration, data protection, and maintainability
	Purchase or build a simulator for	
	• Live simulations with professional actors	• Select suitable actors in terms of sex, age, and physical condition and train them in accordance with guidelines (Lewis et al. 2017). Also select and train confederates systematically (Nestel et al. 2014)
	• Digital simulations	• Purchase or develop a computer simulator that can be integrated within or complement the learning and testing environment. Fidelity of the simulation scenario and interaction possibilities are crucial (Crawford 1966; Gaba 2004; Huwendiek et al. 2009; Meller 1997)

Table 4 (continued)

Phase	Activities and subactivities	Recommendations and important literature
	• Role-play simulations	• Set the requirements for role-players and confederates, organize props and facilities (Khan et al. 2013a). Think about preparing participants well for role plays within a short amount of time. Support measures such as tablets or cheat sheets can deliver cues for the case vignettes (Joyner and Young 2006). Check the amount of information contained in the created materials of the different roles and set it deliberately
	Write case vignettes	Content experts should write up the case vignettes based on a blueprint specifying critical features of all cases. They should engage in regular exchange with researchers and consultants to ensure that the content is suitable for assessment and learning purposes and should ask students whether the cases' relevancy and difficulty are appropriate
	Create briefings	Create technical or content briefings that familiarize learners with the situation. These can take the form of a static familiarization (e.g., a video) or an interactive test case to practice
	Create debriefings	Create debriefings that focus on the learning process or outcome. They can be static or adaptive. Consider the standards for debriefing (INACSL Standards Committee 2016a)
	Design instruments that assess knowledge and competences	Design instruments based on the literature; focus pre- and posttests on the intervention content. Choose test format (e.g., MC questions), scoring, and number deliberately. See standards for participant evaluation (INACSL Standards Committee 2016b)
	Construct educational interventions	Determine the specific content of the intervention, the scope of the learning process the intervention targets, and the timing and frequency
	Develop a measurement procedure	Decide which variables have to be assessed. Determine the most sensible frequency and timing of assessment
	Develop an experimenter's guide	Include in this document specifications on participant allocation, all verbal instructions provided by the experimenter, and standardization of the simulation environment

Table 4 (continued)

Phase	Activities and subactivities	Recommendations and important literature
Evaluation and reflection subphase	Carry out expert workshops	Use expert workshops to evaluate the simulation scenarios, case vignettes, debriefings, knowledge tests, and instruments to assess competences. For documentation, it can be sensible to film the experts' discussion and/or record data in a standardized way
	Conduct pilot studies	To evaluate simulation components, pilot studies can help to check whether log data is recorded correctly, usability is decent, and the intervention seems to be effective. To evaluate research materials, think-aloud can be used to tap cognitive processes
	Reflect on the research and design process	Discuss whether the project fulfilled its own and the program's requirements
Conducting empirical investigations	Prepare the developed materials	Book rooms punctually, offer all involved persons time and support to prepare diligently
	Recruit the sample	Curriculum-integration, peer-recruitment and financial compensation boost participation
	Carry out and document a study using	
	• Lab-based data collection	• Place all requisites and cameras in a tested and standardized way (Cheng et al. 2014b). Check the internet connection, video capture, as well as audio and video quality on used computers (Khan et al. 2011). Create regularly backups of local data
	• Web-based data collection	• Communicate hardware and software requirements to participants. Offering more guidance, use web conference tools, or email to communicate with participants. Offering less guidance, provide detailed familiarizations, and debriefings in the web environment
	• Live simulations with professional actors	• Actors and staff used in live simulations can fall sick. Thus, it can be sensible to have an extra person who can fill in. Brief actors before each date for their role and give them feedback after each date. Provide enough breaks for the staff during the study to prevent fatigue. For more recommendations, see Khan et al. (2013b)
	• Digital simulations	• Digital simulations can include concealed mistakes and lacks of standardization. Check thoroughly for such flaws before the study starts. Moreover, participants can encounter technical problems working on digital simulations (e.g., log-in issues or problems in video streaming). Provide support to these participants. Inform your technical partners about important dates of the study to avoid maintenance jobs
	• Role-play simulations	• Familiarize role-players in detail with their role and explain the materials at hand. Help participants who need support to perform their role in a standardized way. Screen for role-players who abandon their role and note down this

Table 4 (continued)

Phase	Activities and subactivities	Recommendations and important literature
Analyzing data and communicating results	Code and prepare data of	
	• Digital simulations	• Statistics programs such as *R* facilitate the transformation of complex log files and detailed process analyses; easier log-file transformations can be accomplished with spreadsheet software. Create a data file in line with recommendations for tidy data (Wickham 2014)
	• Role-play and live simulations	• Design coding schemes for video recordings of behaviour based on conceptual frameworks. Test coding schemes on a small data sample. Then, have two trained raters independently code the data and report inter-coder agreement
	Analyze data	Use a statistical software such as *R* or *R Markdown* that allows for the creation of reproducible analyses by saving scripts. Check assumptions for all statistical analyses performed
	Create scientific publications	Write empirical articles based on project data; these should follow reporting guidelines (Cheng et al. 2016). Exchange information between research projects on how certain operationalizations are implemented, and data is interpreted. Also, research projects should join forces to write joint articles and submit symposia. Presenting work at conferences is important to receive feedback, train presentation skills, and build professional networks
	Increase visibility of findings	Advertise scientific publications on social media and in blogs. Disseminate research project findings also using popular media formats (e.g., on Twitter or ResearchGate)

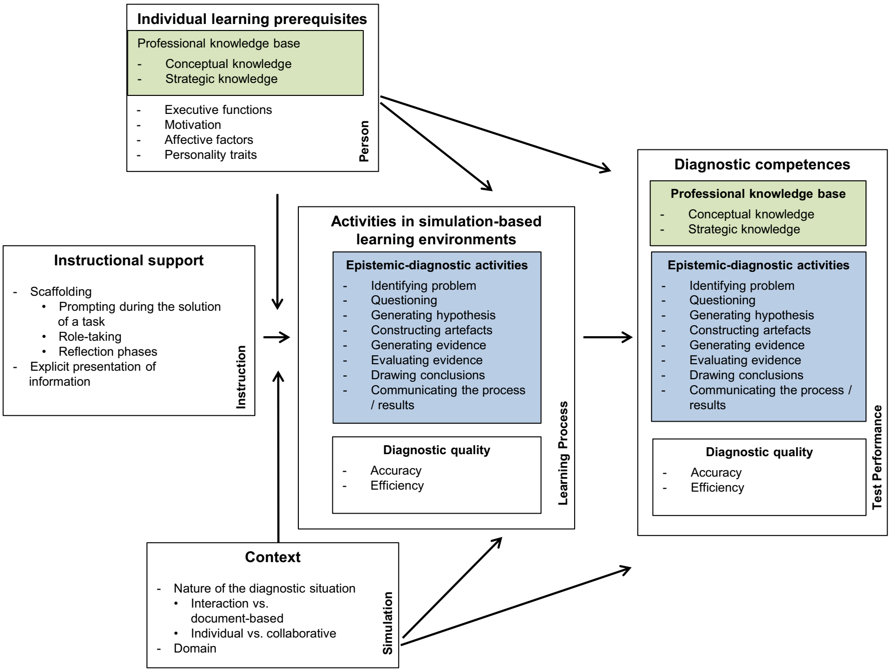


Fig. 2 A framework for facilitating diagnostic competences with simulations (Heitzmann et al. 2019)

consultants, whether it is more sensible to purchase and customize or to newly develop simulators and simulations.

Illustration

The research committee for COSIMA consists of 12 professors from the fields of medical education, teacher education, and educational psychology. Several postdocs and Ph.D. students also carry out tasks for the research committee. The collaborators who formed the research committee for the research program COSIMA drafted a proposal concerning the use of simulations to facilitate diagnostic competences, including a literature review, a preliminary conceptual model (see Fig. 2), and a mentoring concept for research projects. After the official start of the research program, the research committee stated common objectives at a kick-off retreat. It also provided background knowledge on conceptualizing and assessing diagnostic competences in a monthly colloquium.

The principal investigators of the research project COSIMA 4 are two professors of medical education and one professor of educational psychology, who are also part of the research committee. Before the official project start, the principal investigators selected conducting a medical interview in an emergency room as a relevant professional situation for the proposal and specified four studies with a cross-sectional experimental design. Moreover, they decided to simulate the medical interview in the four studies with live simulations with professional actors, digital simulations, and role-play simulations. They also selected different types of reflection phases and roles as the form of instructional support to investigate. After the official project start, a Ph.D. student in learning sciences and a board-certified general practitioner joined the team. An external IT company was contracted to produce the learning and testing environments. Next, the Ph.D. student and general practitioner acquainted themselves with the literature on instructional support, simulations, and history-taking and conducted a site visit to a simulation center.

At both levels, this phase began about one year before the submission of the grant proposals. It was put on hold during the evaluation of the proposals and was completed approximately half a year after the official start of the research program.

Phase 2: constructing and adapting simulations and research materials

Research committee level

In this phase, the research committee creates research and design requirements for the research projects, including operationalizations and specifications for the research materials and simulation components (see Table 3). The fulfillment of these requirements in the projects is monitored (Cheng et al. 2017) throughout this phase, and researchers and designers are guided in fulfilling the requirements. In addition, the research committee conveys a common knowledge base (Sonnenwald 2007) on the content and the design of research materials and simulation components. Apart from these tasks, the research committee may support the research projects as needed with tasks such as ethical approval, legal issues with external contractors, and conflict resolution (Cheng et al. 2017; Hall et al. 2012).

Research project level

Simulations and research materials are constructed and adapted in the research projects following two phases of the generic model for EDR (McKenney and Reeves 2012): 1) design and construction, and 2) evaluation and reflection.

Design and construction

In this subordinate phase, team members generate and assess possible resolutions, specify requirements, and develop products (McKenney and Reeves 2012). The activities performed by each team member depend considerably on their role and the developed research materials and simulation components (see Tables 1 and 2 for definitions and Table 4 for specific recommendations). However, we can summarize that consultants and researchers mainly decide on a learning and testing environment, design instruments that assess knowledge and competences, develop a measurement procedure, and create an experimenter's guide. Designers primarily focus on building a simulator and a learning and testing environment. Content experts focus on creating simulation scenarios, case vignettes as well as briefings and debriefings. As one of the last steps, the developed research materials and simulation components are embedded in the learning and testing environment according to the experimental procedure.

Evaluation and reflection

This subordinate phase consists of a formative evaluation of prototypes of the developed research materials and simulation components through expert workshops and pilot studies, and a deliberate reflection on the research and design process (McKenney and Reeves 2012). Expert workshops typically evaluate the authenticity and difficulty of the developed simulation scenarios and case vignettes. Moreover, expert workshops can scrutinize the content and accuracy of expert solutions presented to participants in debriefings. Pilot studies mainly provide an assessment of the measurement properties of the instruments assessing knowledge and competences and evaluate the usability and fidelity of the simulation. Reflecting on the research and development process, team members evaluate individual work and collaboration, as well as compliance with the milestones of the research project and research program.

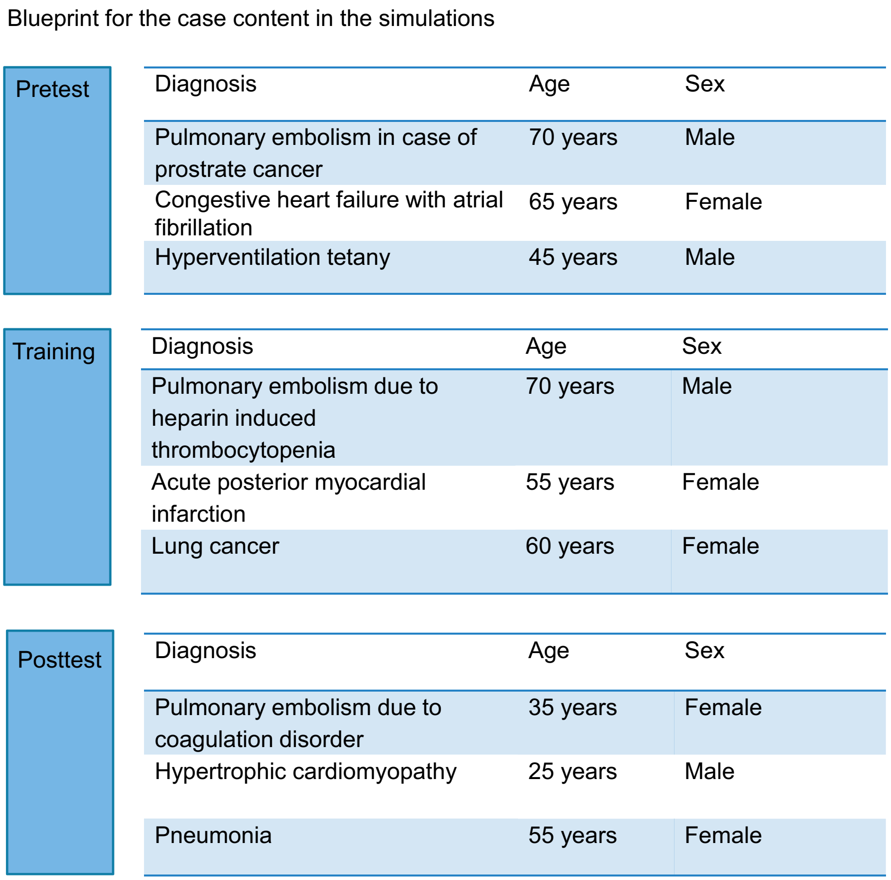
Interaction between the subordinate phases

As various research materials and simulation components are created during the second phase, the two subordinate phases may be iteratively repeated several times. Each of these subordinate phases may focus on one piece of research material or simulation component or several simultaneously. Therefore, multiple cycles may need to be completed before all research materials and simulation components are fully developed and positively evaluated. If the evaluation of all required research materials and simulation components is successful, an empirical investigation can be conducted.

Illustration

In this phase, the research committee of COSIMA guided the development of a common test battery for all research projects. This test battery included motivational and cognitive scales and tests. Moreover, the research committee created a measurement procedure for all research projects that specified the number and timing of measurements for common variables. Throughout this phase, the research committee conveyed knowledge on simulation components and research materials (e.g., on the structure and timing of reflection phases) and offered guidance. In addition to these tasks, the research committee filed for ethical approval in cooperation with the projects.

The research project COSIMA 4 repeated this phase twice. In the first design and construction cycle, prototypes of the digital simulation and live simulation were created. Both prototypes involved only two case vignettes and were created with student assistants as actors. Moreover, knowledge tests on dyspnea were created. These materials were integrated into the selected learning and testing environment CASUS (Instruct 2018). Afterwards, a pilot study primarily investigated the usability of the simulation prototypes. Based on the results of the pilot study, a second design and construction cycle took place. In this cycle, a professional programmer was contracted with improving the created digital simulation. The professional programmer developed a simplified version of the digital simulation with fast-streaming video clips. To improve the live simulations, an acting coach was hired. The acting coach supported the project in preparing and recruiting professional standardized patients. Subsequently, nine case vignettes for the two types of simulations were created (see the blueprint in Fig. 3), and an expert workshop evaluated their accuracy


Blueprint for the case content in the simulations

Pretest

Diagnosis	Age	Sex
Pulmonary embolism in case of prostrate cancer	70 years	Male
Congestive heart failure with atrial fibrillation	65 years	Female
Hyperventilation tetany	45 years	Male

Training

Diagnosis	Age	Sex
Pulmonary embolism due to heparin induced thrombocytopenia	70 years	Male
Acute posterior myocardial infarction	55 years	Female
Lung cancer	60 years	Female

Posttest

Diagnosis	Age	Sex
Pulmonary embolism due to coagulation disorder	35 years	Female
Hypertrophic cardiomyopathy	25 years	Male
Pneumonia	55 years	Female

Fig. 3 Blueprint for the case content in the simulations

and difficulty. After some revisions, the quality of the research materials and simulation components was evaluated positively, and the next phase began.

This phase was completed by the research committee one year after the official program start. Research project COSIMA 4 finished this phase approximately one and a half years after the official program start.

Phase 3: conducting empirical investigations

Research committee level

In this phase, the research committee specifies final targets for the empirical investigations within the research projects (e.g., regarding sample size and composition) in alignment with the research program's initial goals and current status. Also, the research committee monitors the standardization across research projects (Cheng et al. 2017). Moreover, the research committee monitors whether the empirical investigations are proceeding without issue and whether issues co-occur in research projects. Some co-occurring issues can be resolved more easily by a superordinate project, including senior members, than by individual research projects. Besides, the research committee provides guidance for the research projects regarding individual problems (Cheng et al. 2017). Knowledge is primarily conveyed on conducting empirical investigations with simulations, recruiting participants, and recording data.

Research project level

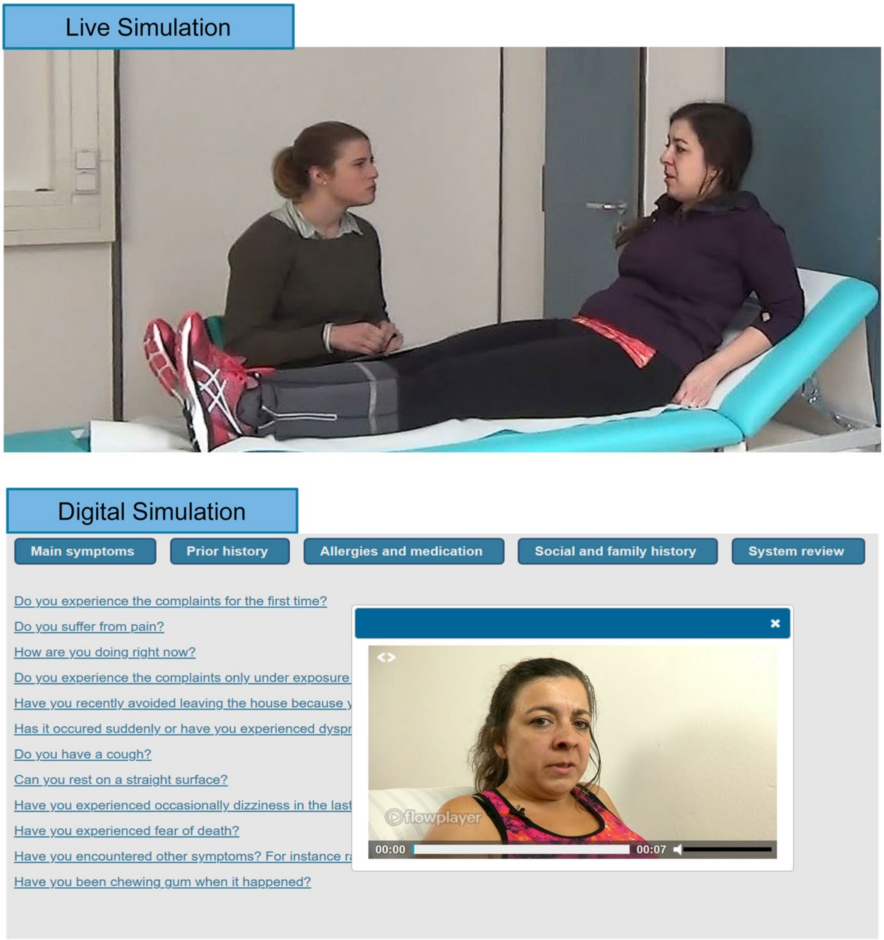
In this phase, researchers prepare the developed materials, recruit the sample of the target population, and carry out and document the empirical study. Regarding the preparation of the developed materials, facilities must be organized in advance, and involved personal must be trained. Regarding recruitment, participants can be recruited on or off campus as well as online. Boosting recruitment, the importance of incentives should also be stressed. Concerning carrying out the study, there are two important types of data collection: lab-based data collection and web-based data collection. Both types of data collection come with particular advantages and disadvantages (see Reips 2000) and we offer recommendations for both in Table 4. Moreover, we believe there are also peculiarities in simulation research that depend on the simulation modality. We provide recommendations for carrying out live simulations with professional actors, digital simulations, and role-play simulations in Table 4. With regards to documentation, funding agencies typically recommend the meticulous documentation of research materials and the empirical investigation as well as open science practices such as sharing data (Earle et al. 2013). Lastly, complex investigations involving multiple simulators, instruments, or many participants have to be organized efficiently, saving costs and minimizing the use of facilities. For example, complex investigations may require protocols that organize the flow of participants. Such protocols can guarantee that participants complete different parts of the investigation punctually and in the correct sequence (for more recommendations on this topic, see Khan et al. 2013a).

Illustration

During the first phase of empirical investigations, the research committee of COSIMA adhered to the initially proposed sample size in all research projects. It conveyed knowledge on organizing simulation studies (e.g., training standardized patients) and recording log data and videos. In addition, the research committee supported research projects when issues with recruitment or facilities (e.g., simulation centers) occurred.

In this phase, the research project COSIMA 4 successfully conducted a full empirical study with $N=86$ medical students (see Fig. 4 for a screenshot of live and digital simulations). Behavior was captured with a camera, and data was recorded with log files recording all activities of participants.

This phase was completed by the research committee for the first time one and a half years after the official program start. The research project COSIMA 4 completed this phase for the first time one year and eight months after the official program start.



Do you experience the complaints for the first time?
Do you suffer from pain?
How are you doing right now?
Do you experience the complaints only under exposure
Have you recently avoided leaving the house because y
Has it occured suddenly or have you experienced dyspr
Do you have a cough?
Can you rest on a straight surface?
Have you experienced occasionally dizziness in the last
Have you experienced fear of death?
Have you encountered other symptoms? For instance r
Have you been chewing gum when it happened?

Fig. 4 Screenshots from the empirical study

Phase 4: analyzing data and communicating results

Research committee level

In this phase, the research committee provides data reporting specifications to all empirical research projects (Tobi and Kampen 2018) and gathers data systematically. When data from several projects has been merged, the research committee can use special methods to analyze large data sets (see recommendations in Table 3). Also, the research committee can convey knowledge on data analysis to the research projects by organizing methods workshops. Above all, the research committee can create in this phase publications based on project data as well as publications that do not rely on project data (see recommendations in Table 3). Moreover, the research committee can develop a strategy that helps the research projects to increase the visibility of findings. Such a strategy should address how publications and key findings are advertised and reported in a comprehensible manner, also understandable to a non-scientific audience. Finally, the research committee also starts to generate first ideas for writing extension grants.

Research project level

This phase frequently starts with coding and preparing data. For digital simulations, log-file data must be transformed, cleaned, and scored. For role-play and live simulations, video recordings of behaviour must be coded with coding schemes. Then, suitable statistical analyses are applied. Results are communicated in such research projects mainly through theoretical and empirical articles and talks at conferences. Moreover, the research projects also make their research more visible to a broader audience by reporting key findings also in popular media formats (e.g., on Twitter or ResearchGate).

Illustration

The research committee of COSIMA created detailed specifications for reporting results in the research projects and compiled data from all seven empirical projects into one file. The research committee has published two conceptual articles so far that do not rely on project data (Förtsch et al. 2018; Heitzmann et al. 2019). In addition, a registered report is in preparation that will evaluate aggregated data from multiple projects using structural equation modeling techniques.

After completing the first study for the research project COSIMA 4, data from the digital simulations' log files was transformed using scripts in the statistical software *R*. The behavior of learners in live simulations was coded by student assistants using a coding scheme. First analyses of the data have been presented in conference talks, and one empirical article on the study has been written and submitted for peer-review.

This phase was completed by the research committee for the first time two years after the official program start. For the research project, this phase ended for the first time two years and three months after the official program start.

Interaction between the phases and subsequent research process

Within the research committee and research projects, the first and second phases usually have to be completed only once and are followed by iterations of the third and fourth phases. If a study requires additional development, the second phase must be repeated. Analyzing data and communicating results in phase 4 for the research committee may begin earlier than for the subordinate research projects (e.g., via the creation of theoretical articles for the projects) and may continue even after all projects have ended. When the research program comes to an end, the research committee and research projects have three options. They can write another proposal for a grant with a new topic, submit a request for an extension, or finish research on this topic.

Illustration

The research committee successfully applied for an extension grant for a second phase of the program together with the research projects after two and a half years. This application process led to a short additional cycle of the first phase. Overlapping with this phase, the research committee worked on publishing findings and was thus in the fourth phase. The research project COSIMA 4 was during that time simultaneously in the third and fourth phases, conducting empirical studies and publishing findings from its first empirical study.

Illustration of roles taken and products created

In the research committee, all professors served as consultants, as they made strategic decisions concerning the research program. All professors also acted as researchers, developing a common theoretical framework and working on publications. Certain professors engaged on this level as content experts, choosing specific topics and contexts for the research program. In the described research project COSIMA 4 , the Ph.D. students mainly served as researcher and instructional designer; the general practitioner acted primarily as content expert. The professors primarily took on the roles of consultants in the research project, guiding the Ph.D. student and general practitioner. The contracted IT companies served as instructional designers by providing and developing the learning and testing environments.

Over the course of the phases, the research committee and research projects created scientific knowledge as well as practical applications and tools. Scientific knowledge includes codified knowledge (such as theoretical and empirical articles) and uncodified knowledge (such as employees' implicit knowledge about creating simulations) within a team (Nonaka 1994). Codified scientific knowledge is mainly created in the fourth phase, while uncodified knowledge develops throughout all phases. Practical applications and tools were primarily developed in the second phase. Particularly intricate applications and tools that were not strictly necessary for the research program were developed in subsidiary projects or will be refined after the end of the research program.

Reasons for the framework's suitability for research programs

Finally, we provide reasons for why the presented framework is particularly suitable to describe research and design in dual-level, multi-project research programs. The main argument for our framework's suitability in this context is its successful modelling of the aforementioned dual-level structure. As we have seen, team members' activities on both levels are rather heterogeneous (see also Tables 3 and 4). Moreover, roles cannot be strictly assigned to certain team members on each level. In contrast, team members took on multiple roles that changed to some extent over the course of the research program. Also, the schedules of the research committee and research projects varied, leading to a situation in which the two levels were in different program/project phases at a given point in time.

Discussion

Summary

The proposed framework elucidates research and design in multi-project research programs empirically investigating simulation-based learning. The framework was derived on the basis of theory and specifies the activities of a research committee and subordinate research projects in four program/project phases in order to create scientific knowledge as well as practical applications and tools. Illustrations from the research program COSIMA exemplified that the basic premises of this framework are applicable. As expected, the program/project phases were completed at different times on each level, and iterated and overlapped considerably. The illustration of research project COSIMA 4 demonstrated that the second phase of the proposed framework represents the simultaneous research and design of simulation components and research materials well.

Implications for theory

Contributing to the EDR literature, the proposed framework complements other descriptive models of simultaneous research and design processes. The proposed framework is briefly compared to the most similar frameworks from this literature that served as models. At the same time, however, the ways in which our framework goes beyond existing models are also clarified. The ILDF (Bannan-Ritland 2003) is a general research and design framework directed at the program level comprising the phases of exploration, enactment, and evaluation. The ILDF applies a rapid prototyping approach and addresses the adoption of created artifacts. In comparison to the proposed framework, the ILDF is more general, does not delineate leadership tasks and seems more geared towards developing artifacts for applied research. The generic model for EDR (McKenney and Reeves 2012) and the osmotic model by Ejersbo et al. (2008) address research and design on a project level and posit the equal status of research and design throughout the project. Due to the equal status of research and design in all phases of these frameworks, the generic model for EDR and the osmotic model may be most suitable to characterize and guide applied research in regular EDR contexts. We believe, however, that our framework is more suitable for research programs that focus on basic research and thus often contain only *one* (iterative) phase in which research and design are equally important. The "compleat" design experiment model (Middleton et al. 2008) integrates the design of an empirical investigation into an educational research process. This model is to some extent similar to our proposed framework because both locate research and design simultaneously in the middle phase of an extensive research process. Contrary to the proposed framework, the "compleat" design experiment focusses on a single research project instead of a research program. Moreover, it describes the development of regular randomized controlled trials instead of research materials and simulations for an empirical investigation of simulation-based learning. In the model from Akkerman et al. (2013), the three epistemic practices of research, design, and educational change are on an equal level over the entire course of a single research project. Consequently, this model seems more suitable than our framework for contexts like translational science in which educational change (i.e., adoption of a created intervention) is crucial. Concluding this analysis of similar models and the described theoretical background, the proposed framework is the only dual-level, multi-project EDR framework. It places a stronger emphasis on research than the other described frameworks, with the exception of the "compleat" design experiment (Middleton et al. 2008). Moreover, the proposed framework is more context- and content-specific than the other EDR models. Concerning the context-specificity, the proposed framework demonstrates that the organizational structure (e.g., team level) can significantly affect EDR and that theories from other fields, such as team science or organizational psychology, can offer valuable insights for representing the context in EDR frameworks. With regard to the content-specificity, our proposed framework is more specific than the other EDR frameworks due to its focus on simulations that assess and facilitate knowledge and competences in medical education and teacher education. In the proposed framework, the challenging need to develop complex content for simulation components and research materials probably led to the equal status of research and design in the second program/project phase. These conclusions on context- and content- specificity show that educational design researchers who create novel EDR frameworks should not hesitate to explore specific contexts and contents that pose unique demands on research and design.

The presented framework also adds to the literature on simulation research and design in medical education and teacher education. In contrast to our framework, the framework by Cheng et al. (2017) focusses on carrying out one empirical study across different simulation centers. Similarly to our framework, Cheng et al.'s (2017) framework includes recommendations for different phases that can serve as a valuable resource for research and design. Other frameworks and models are more specific than our framework, and focus on specific fields such as pediatrics (Cheng et al. 2014a), live simulations (Khan et al. 2013a; Sturpe and Schaivone 2014), and digital simulations (Posel et al. 2015; Zary et al. 2006). Our framework extends this literature by providing an overview of all typical research materials and simulation components that have to be designed and developed and recommendations for this process. By focusing on simulations assessing knowledge and competences, our framework remains sufficiently specific, while being still applicable to many other contexts.

Implications for practice

Leaders of research programs, as well as entire research committees, may gain from the presented framework for conducting such programs. The overview of and recommendations for activities outlined for the research committee level (see Table 3) provide valuable ideas that can be followed and included in grant proposals, which increasingly comprise a section on project coordination. Leaders of research programs and entire research committees can also use the presented framework to track in what phases research programs and projects currently are, what specific research materials and simulation components have been produced, what activities are taking place and which role different team members are playing. These points may improve the management and coordination of research projects, help to conduct state-of-the-art research, and facilitate standardization.

Members of research projects can follow the recommendations for activities in the described project phases and read up on the literature resources provided for certain aspects (see Table 4). They can also use the presented framework to monitor in what phases their research projects currently are, what specific research materials and simulation components have been developed, what activities are taking place, and what role they are playing. These points should support members of research projects in developing high-quality simulation components and research materials, help to conduct state-of-the-art research, and keep the project within its timeline and budget.

Limitations of the framework and directions for future research

Of course, the presented framework is not without limitations that require consideration. One limitation of the presented framework results from the employed methodology. The framework was derived from theory, but illustrated through a research program in which the authors are active members. At first glance, this methodology may seem less scientific to educational researchers, who primarily conduct experiments in highly-standardized laboratory settings (Collins et al. 2004). However, as EDR "focuses on understanding the messiness of real-world practice" (Barab and Squire 2004, p. 3), such a methodology is required for the development of new context-specific frameworks and has been employed in the development of other EDR frameworks (e.g., Bannan-Ritland 2003). Another limitation of the presented framework is that only one illustration of a research committee and one illustration of a research project from the domain of medical education were provided in the text. Even though this small number of illustrations was necessary to convey our framework in a comprehensible way, it also impairs to some extent the framework's generalizability to other domains, notably teacher education. In response to these two points of criticism, we have provided reasons for why our framework depicts this type of research program well and specified commonalities between research and design in medical education and teacher education.

The described framework touches upon two topics that seem especially promising for future research independent of the proposed framework. One interesting topic for future EDR frameworks is team science. The organizational structures and collaborative processes in multi-team research and design programs seem to have not yet been sufficiently addressed in EDR frameworks (compare the frameworks presented in McKenney and Reeves 2012). For instance, new EDR frameworks could put a stronger emphasis on the EDR process and integrate models of collaborative problem-solving (see Graesser et al. 2018). Another exciting avenue of research concerns EDR methodology. With the rise of new technologies such as online collaboration software, research and design processes and the created products are increasingly stored electronically. This development will allow educational design researchers to validate EDR frameworks and investigate research and design processes using electronic data without noticeable intrusions. This validation of EDR frameworks with electronic data would respond to a long-standing request for the EDR literature (Kelly 2006).

With regard to the proposed framework, it would be interesting to find out whether the described phases and activities also occur in research programs from other domains. There are other domains in which training knowledge and competences with simulations could be investigated fruitfully within a research program. For instance, simulations are already used frequently for the purposes described above in nursing education (Cant and Cooper 2017) and should become more and more popular in foreign language learning in the future (Blyth 2018). Moreover, it could be investigated whether the presented framework can be applied successfully in research programs that develop complex artifacts other than simulations, such as serious games. Serious games refer to the creation of complex and immersive virtual worlds that resemble digital simulations. Thus, their design process could follow rather similar principles to that of simulations (Kirkley et al. 2007). Research programs on serious games could take away from our model recommendations for implementing assessment and training, for example. Apart from this, it remains an open question whether the presented framework is applicable in different organizational structures. The framework has been illustrated with one interdisciplinary research program that was conducted at multiple research groups at several universities. Larger research centers, for instance, also have a relatively strong hierarchical structure but nevertheless coordinate several full research programs and their associated research projects on a superordinate level.

Conclusions

We have presented a framework for research and design in research programs consisting of one research committee and multiple research projects. This framework is particularly suitable for research programs that focus on simulation-based learning but could potentially also be used in other research and design settings that comprise the design of complex scientific artifacts. On a theoretical level, the framework contributes primarily to the literature on EDR by offering a unique dual-level perspective that delineates relevant phases, activities, roles, and products. Moreover, on a practical level, the framework may guide the research and design process of leaders and members of research programs. We hope that the checklists provided in this article are a valuable resource for this aim.

Funding Open Access funding enabled and organized by Projekt DEAL. The research presented in this article was funded by the German Research Association (Deutsche Forschungsgemeinschaft, DFG), project number FOR2385, and the German Federal Ministry of Education and Research (BMBF) under the references 16DHL1039 and 16DHL1040.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

Akkerman, S. F., Bronkhorst, L. H., & Zitter, I. (2013). The complexity of educational design research. *Quality & Quantity, 47*(1), 421–439. https://doi.org/10.1007/s11135-011-9527-9.

Badiee, F., & Kaufman, D. (2015). Design evaluation of a simulation for teacher education. *SAGE Open, 5*(2), 215824401559245. https://doi.org/10.1177/2158244015592454.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21–24. https://doi.org/10.3102/0013189X032001021.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1.

Barrows, H. S., & Abrahamson, S. (1964). The programmed patient: A technique for appraising student performance in clinical neurology. *Academic Medicine, 39*(8), 802–805.

Beck, S. J., Meinecke, A. L., Matsuyama, Y., & Lee, C.-C. (2017). Initiating and maintaining collaborations and facilitating understanding in interdisciplinary group research. *Small Group Research, 48*(5), 532–543. https://doi.org/10.1177/1046496417721746.

Benishek, L. E., Lazzara, E. H., Gaught, W. L., Arcaro, L. L., Okuda, Y., & Salas, E. (2015). The template of events for applied and critical healthcare simulation (TEACH Sim): A tool for systematic simulation scenario design. *Simulation in Healthcare, 10*(1), 21–30. https://doi.org/10.1097/SIH.0000000000000058.

Bennett, M. L., & Gadlin, H. (2012). Collaboration and team science: From theory to practice. *Journal of Investigative Medicine, 60*(5), 768–775. https://doi.org/10.2310/JIM.0b013e318250871d.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Für Psychologie, 223,* 3–13. https://doi.org/10.1027/2151-2604/a000194.

Blyth, C. (2018). Immersive technologies and language learning. *Foreign Language Annals, 51*(1), 225–232. https://doi.org/10.1111/flan.12327.

Bozeman, B., & Boardman, C. (2014). *Research collaboration and team science: A state-of-the-art review and agenda. Springer Briefs in Entrepreneurship and Innovation*. Cham: Springer. https://doi.org/10.1007/978-3-319-06468-0.

Butler, D. (2008). Translational research: Crossing the valley of death. *Nature, 453*(7197), 840–842. https://doi.org/10.1038/453840a.

Cant, R. P., & Cooper, S. J. (2017). Use of simulation-based learning in undergraduate nurse education: An umbrella systematic review. *Nurse Education Today, 49,* 63–71. https://doi.org/10.1016/j.nedt.2016.11.015.

CASUS (Version 2.0) [Computer software] (2018). Retrieved from https://www.instruct.eu/.

Charlin, B., Gagnon, R., Lubarsky, S., Lambert, C., Meterissian, S., Chalk, C., … van der Vleuten, C. (2010). Assessment in the context of uncertainty using the script concordance test: More meaning for scores. *Teaching and Learning in Medicine*, *22*(3), 180–186. https://doi.org/10.1080/10401334.2010.488197.

Cheng, A., Auerbach, M., Hunt, E. A., Chang, T. P., Pusic, M., Nadkarni, V., & Kessler, D. (2014). Designing and conducting simulation-based research. *Pediatrics, 133*(6), 1091–1101. https://doi.org/10.1542/peds.2013-3267.

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: A systematic review and meta-analysis. *Medical Education, 48*(7), 657–666. https://doi.org/10.1111/medu.12432.

Cheng, A., Kessler, D., Mackinnon, R., Chang, T. P., Nadkarni, V. M., Hunt, E. A., … Auerbach, M. (2016). Reporting guidelines for health care simulation research: Extensions to the CONSORT and STROBE statements. *Advances in Simulation*. https://doi.org/10.1186/s41077-016-0025-y.

Cheng, A., Kessler, D., Mackinnon, R., Chang, T. P., Nadkarni, V. M., Hunt, E. A., … Auerbach, M. (2017). Conducting multicenter research in healthcare simulation: Lessons learned from the INSPIRE network. *Advances in Simulation*, *2*, 6. https://doi.org/10.1186/s41077-017-0039-0.

Christensen, R., Knezek, G., Wood, T. T., & Gibson, D. (2011). SimSchool: An online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technology, 6*(2), 201–220. https://doi.org/10.1504/IJLT.2011.042649.

Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher, 45*(1), 48–54. https://doi.org/10.3102/0013189X16631750.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1), 15–42.

Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., … Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Medical Teacher*, *35*(1), e867–e898. https://doi.org/10.3109/0142159X.2012.714886.

Crawford, M. P. (1966). Dimensions of simulation. *American Psychologist, 21*(8), 788–796. https://doi.org/10.1037/h0023974.

De Coninck, K., Valcke, M., Ophalvens, I., & Vanderlinde, R. (2019). Bridging the theory-practice gap in teacher education: The design and construction of simulation-based learning environments. In K. Hellmann, J. Kreutz, M. Schwichow, & K. Zaki (Eds.), *Kohärenz in der Lehrerbildung: Theorien, Modelle und empirische Befunde* (pp. 263–280). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-23940-4_17.

De Jong, T. (1991). Learning and instruction with computer simulations. *Education and Computing, 6*(3–4), 217–229.

DeCostanza, A., DiRosa, G., Jiménez-Rodríguez, M., & Cianciolo, A. (2014). No mission too difficult: Army units within exponentially complex multiteam systems. Pushing the boundaries: Multiteam systems in research and practice. In M. L. Shuffler, R. Rico, & E. Salas (Eds.), *Research on managing groups and teams* (Vol. 16, pp. 61–76). Bingley: Emerald.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Dotger, B. H., Dotger, S. C., & Maher, M. J. (2010). From Medicine to Teaching: The Evolution of the Simulated Interaction Model. *Innovative Higher Education, 35*(3), 129–141. https://doi.org/10.1007/s10755-009-9128-x.

Ejersbo, L. R., Engelhardt, R., Frølunde, L., Hanghøj, T., Magnussen, R., & Misfeldt, M. (2008). Balancing product design and theoretical insights. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.), *Handbook of* *design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching* (pp. 149–164). New York: Routledge.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare, 2*(2), 115–125.

Ferry, B., & Kervin, L. (2007). Developing an online classroom simulation to support a pre-service teacher education program. In D. A. Gibson, C. Aldrick, & M. Prensky (Eds.), *Games and simulations in online learning: Research and development frameworks* (pp. 189–205). Hershey: Information Science Publishing.

Finley, F. N., & Pocoví, M. C. (2000). Considering the scientific method of inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 47–62). Washington, D.C.: American Association for the Advancement of Science.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., … Fischer, M. R. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, *5*, 28–45. https://doi.org/10.14786/flr.v2i3.96.

Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M. R., Girwidz, R., Obersteiner, A., … Neuhaus, B. J. (2018). Systematizing professional knowledge of medical doctors and teachers: Development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, *8*(4), 207. https://doi.org/10.3390/educsci8040207.

Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and Safety in Health Care, 13*(Suppl 1), i2–i10. https://doi.org/10.1136/qshc.2004.009878.

Gagné, R. M. (2007). *Principles of instructional design* (5th ed.). Belmont: Thomson/Wadsworth.

Gartmeier, M., Bauer, J., Fischer, M. R., Hoppe-Seyler, T., Karsten, G., Kiessling, C., … Prenzel, M. (2015). Fostering professional communication skills of future physicians and teachers: Effects of e-learning with video cases and role-play. *Instructional Science*, *43*(4), 443–462. https://doi.org/10.1007/s11251-014-9341-6.

Gitlin, L. N., & Lyons, K. J. (2013). *Successful grant writing: Strategies for health and human service professionals*. New York: Springer.

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

Gray, B. (2008). Enhancing transdisciplinary research through collaborative leadership. *American Journal of Preventive Medicine, 35*(2), S124–S132. https://doi.org/10.1016/j.amepre.2008.03.037.

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.

Hall, K. L., Vogel, A. L., Stipelman, B., Stokols, D., Morgan, G., & Gehlert, S. (2012). A four-phase model of transdisciplinary team-based research: Goals, team processes, and strategies. *Translational Behavioral Medicine, 2*(4), 415–430. https://doi.org/10.1007/s13142-012-0167-y.

Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective structured examination. *British Medical Journal, 1*(5955), 447–451. https://doi.org/10.1136/bmj.1.5955.447.

Heitzmann, N., Fischer, M. R., & Fischer, F. (2017). Towards more systematic and better theorised research on simulations. *Medical Education, 51*(2), 129–131. https://doi.org/10.1111/medu.13239.

Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M. R., … Fischer, F. (2019). Facilitating diagnostic competences in simulations in higher education: A framework and a research agenda. *Frontline Learning Research, 4*, 1–24. https://doi.org/10.14786/flr.v7i4.384.

Hirumi, A., Johnson, T., Reyes, R. J., Lok, B., Johnsen, K., Rivera-Gutierrez, D. J., … Kleinsmith, A. (2016). Advancing virtual patient simulations through design research and interPLAY: Part II—integration and field test. *Educational Technology Research and Development*, *64*(6), 1301–1335. https://doi.org/10.1007/s11423-016-9461-6.

Hirumi, A., Kleinsmith, A., Johnsen, K., Kubovec, S., Eakins, M., Bogert, K., … Cendan, J. (2016). Advancing virtual patient simulations through design research and interPLAY: Part I: design and development. *Educational Technology Research and Development*, *64*(4), 763–785. https://doi.org/10.1007/s11423-016-9429-6.

Huffman, J. L., McNeil, G., Bismilla, Z., & Lai, A. (2016). Essentials of scenario building for simulation-based education. In V. J. Grant & A. Cheng (Eds.), *Comprehensive healthcare simulation: Pediatrics* (pp. 19–29). Cham: Springer. https://doi.org/10.1007/978-3-319-24187-6_2.

Huwendiek, S., de Leng, B. A., Zary, N., Fischer, M. R., Ruiz, J. G., & Ellaway, R. (2009). Towards a typology of virtual patients. *Medical Teacher, 31*(8), 743–748.

INACSL Standards Committee. (2016a). INACSL standards of best practice: SimulationSM debriefing. *Clinical Simulation in Nursing, 12,* S21–S25. https://doi.org/10.1016/j.ecns.2016.09.008.

INACSL Standards Committee. (2016b). INACSL standards of best practice: SimulationSM participant evaluation. *Clinical Simulation in Nursing, 12,* S26–S29. https://doi.org/10.1016/j.ecns.2016.09.009.

Jensen, S., Kushniruk, A. W., & Nøhr, C. (2015). Clinical simulation: A method for development and evaluation of clinical information systems. *Journal of Biomedical Informatics, 54,* 65–76. https://doi.org/10.1016/j.jbi.2015.02.002.

Jones, F., Passos-Neto, C. E., & Braghiroli, O. F. (2015). Simulation in medical education: Brief history and methodology. *Principles and Practice of Clinical Research, 1*(2), 56–63.

Joyner, B., & Young, L. (2006). Teaching medical students using role play: Twelve tips for successful role plays. *Medical Teacher, 28*(3), 225–229. https://doi.org/10.1080/01421590600711252.

Kastner, M., & Stangla, B. (2011). Multiple choice and constructed response tests: Do test format and scoring matter. *Procedia - Social and Behavioral Sciences, 12,* 263–273. https://doi.org/10.1016/j.sbspro.2011.02.035.

Kaufman, D., & Ireland, A. (2016). Enhancing teacher education with simulations. *TechTrends, 60*(3), 260–267. https://doi.org/10.1007/s11528-016-0049-0.

Kelly, A. E. (2006). Quality criteria for design research. In J. van den Akker, K. Gavemeijer, S. E. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 166–184). Abingdon: Routledge.

Khan, K. Z., Gaunt, K., Ramachandran, S., & Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE): Amee Guide No. 81. Part II: Organisation & administration. *Medical Teacher*, *35*(9), e1447-e1463. https://doi.org/10.3109/0142159X.2013.818635.

Khan, K. Z., Ramachandran, S., Gaunt, K., & Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE): Amee Guide No. 81. Part I: An historical and theoretical perspective. *Medical Teacher, 35*(9), e1437–e1446. https://doi.org/10.3109/0142159X.2013.818634.

Khan, K. Z., Tolhurst-Cleaver, S., White, S., & Simpson, W. (2011). *Simulation in healthcare education. Building a simulation programme: A practical guide*. *AMEE guides: Vol. 2*. Dundee: Association for Medical Education in Europe.

Kirkley, J., Kirkley, S., & Heneghan, J. (2007). Building bridges between serious game design and instructional design: A blueprint for now and the future. In B. E. Shelton, Wiley, & D. A. (Eds.), *The design and use of simulation computer games in education* (pp. 61–83). Rotterdam: Sense Publishers.

Koivisto, J.-M., Hannula, L., Bøje, R. B., Prescott, S., Bland, A., Rekola, L., & Haho, P. (2018). Design-based research in designing the model for educating simulation facilitators. *Nurse Education in Practice, 29,* 206–211. https://doi.org/10.1016/j.nepr.2018.02.002.

Lamé, G., & Dixon-Woods, M. (2018). Using clinical simulation to study how to improve quality and safety in healthcare. *BMJ Simulation and Technology Enhanced Learning*. https://doi.org/10.1136/bmjstel-2018-000370.

Lee, W. W. (1994). Subject matter experts and instructional designers: Making distinctions. *Performance + Instruction*, *33*(8), 23–25. https://doi.org/10.1002/pfi.4160330807.

Lewis, K. L., Bohnert, C. A., Gammon, W. L., Hölzer, H., Lyman, L., Smith, C., … Gliva-McConvey, G. (2017). The association of standardized patient educators (ASPE) standards of best practice (SOBP). *Advances in Simulation*, *2*, 10. https://doi.org/10.1186/s41077-017-0043-4.

McKenney, S. E., & Brand-Gruwel, S. (2018). Roles and competencies of educational design researchers: One framework and seven guidelines. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology* (Vol. 41, pp. 1–26). Cham: Springer. https://doi.org/10.1007/978-3-319-17727-4_123-1.

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

Meller, G. (1997). A typology of simulators for medical education. *Journal of Digital Imaging, 10*(3), 194–196.

Middleton, J., Gorard, S., Taylor, C., & Bannan-Ritland, B. (2008). The "compleat" design experiment: From soup to nuts. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching* (pp. 21–46). New York: Routledge.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63–S67.

Nestel, D., Mobley, B. L., Hunt, E. A., & Eppich, W. J. (2014). Confederates in health care simulations: Not as simple as it seems. *Clinical Simulation in Nursing, 10*(12), 611–616. https://doi.org/10.1016/j.ecns.2014.09.007.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37. https://doi.org/10.1287/orsc.5.1.14.

Ormel, B. J. B., Pareja Roblin, N. N., McKenney, S. E., Voogt, J. M., & Pieters, J. M. (2012). Research–practice interactions as reported in recent design studies: Still promising, still hazy.

Educational Technology Research and Development, 60(6), 967–986. https://doi.org/10.1007/s11423-012-9261-6.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*(3), 293–316.

Posel, N., Mcgee, J. B., & Fleiszer, D. M. (2015). Twelve tips to support the development of clinical reasoning skills using virtual patient cases. *Medical Teacher, 37*(9), 813–818. https://doi.org/10.3109/0142159X.2014.993951.

Reiff, R., Harwood, W. S., & Phillipson, T. (2002). A scientific method based upon research scientists' conceptions of scientific inquiry. In P. A. Rubba, J. A. Rye, W. J. DiBiase, & B. A. Crawford (Chairs), *Proceedings of the Annual International Conference of the Association for the Education of Teachers in Science, 2002,* Charlotte, North Carolina.

Reips, U. D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet*. San Diego, CA: Acadmic Press. https://doi.org/10.5167/uzh-19760.

Rothman, A. I., Blackmore, D., Dauphinee, W. D., & Reznick, R. (1996). The use of global ratings in OSCE station scores. *Advances in Health Sciences Education, 1*(3), 215–219. https://doi.org/10.1007/BF00162918.

Rudolph, J. W., Raemer, D. B., & Simon, R. (2014). Establishing a safe container for learning in simulation: The role of the presimulation briefing. *Simulation in Healthcare, 9*(6), 339–349.

Ryall, T., Judd, B. K., & Gordon, C. J. (2016). Simulation-based assessments in health professional education: A systematic review. *Journal of Multidisciplinary Healthcare, 9,* 69–82. https://doi.org/10.2147/JMDH.S92695.

Shuffler, M. L., Jiménez-Rodríguez, M., & Kramer, W. S. (2015). The science of multiteam systems: A review and future research agenda. *Small Group Research, 46*(6), 659–699. https://doi.org/10.1177/1046496415603455.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23.

Simpson, M. A. (1985). How to use role-play in medical teaching. *Medical Teacher, 7*(1), 75–82.

Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology, 41*(1), 643–681. https://doi.org/10.1002/aris.2007.1440410121.

Sturpe, D. A., & Schaivone, K. A. (2014). A primer for objective structured teaching exercises. *American Journal of Pharmaceutical Education, 78*(5), 104. https://doi.org/10.5688/ajpe785104.

Tobi, H., & Kampen, J. K. (2018). Research design: The methodology for interdisciplinary research framework. *Quality & Quantity, 52*(3), 1209–1225. https://doi.org/10.1007/s11135-017-0513-8.

Wickham, H. (2014). Tidy data. *Journal of Statistical Software, 59*(10), 1–23.

Zary, N., Johnson, G., Boberg, J., & Fors, U. G. H. (2006). Development, implementation and pilot evaluation of a web-based virtual patient case simulation environment-Web-SP. *BMC Medical Education, 6,* 10. https://doi.org/10.1186/1472-6920-6-10.

Ziv, A., Small, S. D., & Wolpe, P. W. (2000). Patient safety and simulation-based medical education. *Medical Teacher, 22*(5), 489–495. https://doi.org/10.1080/01421590050110777.

Maximilian C. Fink works as a research associate at the Institute for Medical Education, University Hospital, LMU Munich. His research interests revolve around the effects of instructional support approaches in simulation-based training, clinical reasoning and educational design research.

Anika Radkowitsch works as a research associate at the Chair of Education and Educational Psychology at LMU Munich. Anika Radkowitsch's main research interests include collaboration and collaborative learning, educational technology, particularly simulation-based learning, and diagnostic competences.

Elisabeth Bauer is a research associate at the Chair of Education and Educational Psychology at LMU Munich. Her major research interests are argumentation in simulation-based learning and adaptive feedback.

Michael Sailer is a postdoctoral scholar at the Chair of Education and Educational Psychology at LMU Munich. He is currently conducting research about game-based learning, with a strong focus on gamification, simulation-based learning, and the use of technology in classrooms and in teacher education.

Jan Kiesewetter is associate professor and research coordinator for the Institute for Medical Education at the University Hospital, LMU Munich. His research interests revolve around individual and cooperative clinical reasoning as well as leadership and argumentation in medicine.

Ralf Schmidmaier is a senior physician and professor for internal medicine at the University Hospital, LMU Munich and is responsible for ward round simulation training and clerkships in internal medicine and deputy director of the Department for Internal Medicine IV.

Matthias Siebeck is a consultant at the Department of Surgery and senior researcher at the Institute for Medical Education at the University Hospital, LMU Munich. Moreover, he is one of the founders of the Center for International Health, an interdisciplinary and international network of excellence at LMU Munich. Matthias Siebeck has published scientific articles in the fields of surgery, global health and teaching with simulations.

Frank Fischer is a full professor of educational science and educational psychology at LMU Munich. He is the speaker of the Munich Center of the Learning Sciences, an interdisciplinary collaboration of more than 30 research groups focusing on advancing research on learning “from cortex to community”. His research focuses on how people learn to engage in scientific reasoning and argumentation, as well as in diagnostic reasoning.

Martin R. Fischer is tenured professor and dean of clinical education at the medical faculty of LMU Munich and the director of the Institute for Medical Education at the University Hospital, LMU Munich. His research interests include clinical reasoning, scientific reasoning, learning with digital media, and interprofessional communication and healthcare.