

Frank Fischer
Ansgar Opitz *Editors*

Learning to Diagnose with Simulations

Examples from Teacher Education and
Medical Education

OPEN ACCESS

 Springer

Learning to Diagnose with Simulations

Frank Fischer • Ansgar Opitz
Editors


Learning to Diagnose with Simulations

Examples from Teacher Education
and Medical Education

 Springer

Editors

Frank Fischer
Chair of Education and Educational
Psychology
Department of Psychology, LMU Munich
Munich, Germany

Ansgar Opitz 
Chair of Education and Educational
Psychology
Department of Psychology, LMU Munich
Munich, Germany



ISBN 978-3-030-89146-6 ISBN 978-3-030-89147-3 (eBook)
<https://doi.org/10.1007/978-3-030-89147-3>

© The Editor(s) (if applicable) and The Author(s) 2022. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

| | | |
|----------|--|-----------|
| 1 | Learning to Diagnose with Simulations: Introduction | 1 |
| | Frank Fischer, Olga Chernikova, and Ansgar Opitz | |
| 2 | A Theoretical Framework for Fostering Diagnostic Competences with Simulations in Higher Education | 5 |
| | Olga Chernikova, Nicole Heitzmann, Ansgar Opitz, Tina Seidel, and Frank Fischer | |
| 3 | Learning to Diagnose Primary Students' Mathematical Competence Levels and Misconceptions in Document-Based Simulations | 17 |
| | Angelika Wildgans-Lang, Sarah Scheuerer, Andreas Obersteiner, Frank Fischer, and Kristina Reiss | |
| 4 | Diagnosing Mathematical Argumentation Skills: A Video-Based Simulation for Pre-Service Teachers | 33 |
| | Elias Codreanu, Sina Huber, Sarah Reinhold, Daniel Sommerhoff, Birgit J. Neuhaus, Ralf Schmidmaier, Stefan Ufer, and Tina Seidel | |
| 5 | Diagnosing 6th Graders' Understanding of Decimal Fractions: Fostering Mathematics Pre-Service Teachers' Diagnostic Competences with Simulated One-on-One Interviews | 49 |
| | Bernhard Marczynski, Larissa J. Kaltefleiter, Matthias Siebeck, Christof Wecker, Kathleen Stürmer, and Stefan Ufer | |
| 6 | Diagnosing the Instructional Quality of Biology Lessons Based on Staged Videos: Developing DiKoBi, A Video-Based Simulation | 63 |
| | Maria Kramer, Julia Stürmer, Christian Förtsch, Tina Seidel, Stefan Ufer, Martin R. Fischer, and Birgit J. Neuhaus | |

| | | |
|-----------|---|------------|
| 7 | Learning to Diagnose Secondary School Students' Scientific Reasoning Skills in Physics and Biology: Video-Based Simulations for Pre-Service Teachers | 83 |
| | Amadeus J. Pickal, Christof Wecker, Birgit J. Neuhaus, and Raimund Girwidz | |
| 8 | Learning to Diagnose Students' Behavioral, Developmental, and Learning Disorders in a Simulation-Based Learning Environment for Pre-Service Teachers | 97 |
| | Elisabeth Bauer, Michael Sailer, Jan Kieseewetter, Claudia Schulz, Iryna Gurevych, Martin R. Fischer, and Frank Fischer | |
| 9 | Live and Video Simulations of Medical History-Taking: Theoretical Background, Design, Development, and Validation of a Learning Environment | 109 |
| | Maximilian C. Fink, Victoria Reitmeier, Matthias Siebeck, Frank Fischer, and Martin R. Fischer | |
| 10 | Diagnosing Collaboratively: A Theoretical Model and a Simulation-Based Learning Environment | 123 |
| | Anika Radkowitzsch, Michael Sailer, Martin R. Fischer, Ralf Schmidmaier, and Frank Fischer | |
| 11 | Conclusions and Outlook: Toward more Systematic Research on the Use of Simulations in Higher Education | 143 |
| | Ansgar Opitz, Martin R. Fischer, Tina Seidel, and Frank Fischer | |
| | Index | 151 |

Chapter 8

Learning to Diagnose Students' Behavioral, Developmental, and Learning Disorders in a Simulation-Based Learning Environment for Pre-Service Teachers



Elisabeth Bauer, Michael Sailer, Jan Kiesewetter, Claudia Schulz, Iryna Gurevych, Martin R. Fischer, and Frank Fischer

This chapter's simulation at a glance

| | |
|------------------------|--|
| Domain | Teacher education |
| Topic | Students' behavioral, developmental, and learning disorders |
| Learner's task | Take on the role of a teacher and gather information about a problematic student to identify if the student may have a clinically relevant disorder and if so, which one it could be |
| Target group | Pre-service teachers for all school tracks in various stages of teacher education |
| Diagnostic mode | Individual diagnosing |
| Sources of information | Documents (students' school assignments, report cards, etc.); reports of the student's in-class and out-of-class behavior; protocols of conversations with other teachers, the student and parents |
| Special features | Use of natural language processing to provide automatic adaptive feedback based on the learners' written explanations of their diagnostic conclusion and processing of the case |

E. Bauer (✉) · M. Sailer · F. Fischer
Chair of Education and Educational Psychology, Department of Psychology, LMU Munich, Munich, Germany
e-mail: Elisabeth.Bauer@psy.lmu.de

J. Kiesewetter · M. R. Fischer
Institute for Medical Education, University Hospital, LMU Munich, Munich, Germany

C. Schulz · I. Gurevych
Ubiquitous Knowledge Processing (UKP) Lab, Department of Computer Science, Technical University Darmstadt, Darmstadt, Germany

8.1 Competence Goals in Higher Education

Contemporary curricula in higher education emphasize the need to facilitate students' competence development. This trend is supported by practitioners and politicians, arguing that work in the digital age requires not merely conceptual knowledge but also the ability to apply it to complex tasks in ill-defined situations (Ananiadou & Claro, 2009). The emphasis on diagnostic competence development in medical and teacher education is one of many examples related to this trend. In accordance with Fischer et al. (2022), we define the action of diagnosing as the goal-oriented collection and interpretation of case-specific or problem-specific information to reduce uncertainty in order to make medical or educational decisions. Thus, diagnostic competences are indicated by the accuracy of the diagnosis, application of professional knowledge (see Förtsch et al., 2018), and performance of appropriate epistemic-diagnostic activities (see Fischer et al., 2014).

Since learning competences is highly complex, support structures are required that guide learners in their learning process (Chernikova et al., 2022; Van Merriënboer et al., 2002). One such support structure is feedback, which has been shown to be one of the main predictors of learning outcomes (Hattie & Timperley, 2007). Individual feedback requires high time investments by higher education instructors, which is why it is often neglected (Nicol, 2010). This is one example of how changing professional requirements affect learning objectives, which in turn affect higher education practices and requirements.

Simultaneously, digitalization has brought about technical innovations that can help to facilitate the adaptation of higher education practices. In recent decades, computer-supported and web-based learning has enabled the widespread usage of a range of instructional methods and measures for learning support. Among these is simulation-based learning (Baek, 2009; Gegenfurtner et al., 2014), which has been shown to be an effective approach for competence development (Berman et al., 2016). There are also attempts to automate learner support in digital learning environments, such as using artificial intelligence for intelligent tutoring systems (Diziol et al., 2010; Naser, 2012). Such intelligent systems are able to adapt automatically to learners' competence level and learning progress by automatically analyzing log data. Novel approaches also automate the analysis of written answers using natural language processing (NLP) methods. These systems are utilized, for example, to analyze lexical, syntactical, rhetorical, and other features of learners' essays to provide feedback on the essays' quality in terms of writing strategy (McNamara et al., 2013). A more detailed automated analysis of writing strategy in combination with the content of written answers was previously unrealizable due to limitations of natural language processing methods (Diziol et al., 2010).

FAMULUS makes progress on this technical challenge with the most recent natural language processing methodology, namely artificial neural networks, to provide automatic adaptive feedback on learners' written text answers while they are engaged in simulation-based learning, in order to foster their diagnostic

competences. The feedback is conceptualized to consider both the strategy and content applied in the text answers. This combination better approximates more advanced levels of feedback. Moreover, FAMULUS is an interdisciplinary project involving the disciplines of teacher and medical education. The current chapter gives an overview of the project's background, goals, learning environment, schedule and open questions referring to the teacher education subproject.

8.2 Teachers' Diagnosing of Their Students' Psychological Problems

As previous chapters have already outlined (see Chernikova et al., 2022; Codreanu et al., 2022), diagnostic competences are a core learning objective in teacher education. Teachers have to diagnose students' performance (Schrader, 2011) and individual prerequisites, such as competence level and motivation (Spinath, 2005). These individual prerequisites also include students' behavioral, developmental, and learning disorders. Such disorders affect around 5% of students (Hölling et al., 2014). Behavioral disorders like ADHD and developmental disorders like specific learning disorders become observable in elementary school or early secondary school at the latest and are therefore relevant for teachers in all types of schools. Often, the symptomatology further evolves as students face increasing performance-related and social challenges in school (Schulte-Körne, 2016). This is why teachers are confronted with students' behavioral, developmental, and learning disorders in their classrooms. They are oftentimes the first professionals who have the opportunity to identify an existing problem and initiate further action (Reinke et al., 2011). Therefore, diagnosing students' psychological problems is not only a relevant aspect of teachers' everyday practice but part of their professional responsibility. When confronted with a problem, teachers need to apply epistemic activities, like generating hypotheses, generating and evaluating evidence for and against these hypotheses, and drawing diagnostic conclusions (see Fischer et al., 2014). In this regard, diagnosing can be decomposed into the application of a diagnostic strategy (see Fischer et al., 2014) and relevant concept knowledge (see Coderre et al., 2003; see Förtsch et al., 2018). One example would be the evaluation of the evidence for "inattention" and "hyperactivity" to draw a conclusion regarding the hypothesis "ADHD". Teachers should be able to identify psychological problems among students and apply a diagnostic strategy and relevant concepts accordingly. Moreover, they need to be able to communicate their diagnoses professionally (see Lawson & Daniel, 2011) e.g., to a school psychologist. This requires combining arguments for and (if applicable) against differential diagnoses to construct a diagnostic argument.

Despite its relevance, students' psychological problems are rarely part of teachers' initial professional education. It has been found that teachers rate their general knowledge about psychological disorders as mediocre at best (Reinke et al.,

2011; Rothi et al., 2008). Consequently, diagnosing students' psychological problems seems to be a particular challenge for teachers (Eklund et al., 2009; Papandrea & Winefield, 2011; Rothi et al., 2008; Trudgen & Lawn, 2011). Aside from students' families, teachers usually possess the broadest information about individual students. Observations in and outside the classroom, documents like assignments and exams, conversations with other teachers, the students themselves, parents or other students can provide meaningful insights. Moreover, teachers can observe their students over the course of at least one school year and therefore gain a developmental perspective on each student. In particular, externalizing disorders like ADHD that manifest considerably in a student's behavior allow teachers to apply a wide range of observational methods and resources. Other disorders that can be identified by teachers are developmental disorders of scholastic skills like dyslexia, since they have a strong impact on a student's performance.

Generally, the literature on teachers' diagnosing of students' psychological problems is sparse. One reason for that might be that the topic is located at the intersection of two professional disciplines, namely teaching and clinical psychology. These two disciplines as well as adjacent professional disciplines offer valuable insights into teachers' diagnosing and how to design a suitable learning environment for pre-service teachers. The following section further elaborates on the interdisciplinary relations concerning teachers' diagnosing of students' psychological problems.

8.3 Interdisciplinary Setting

The central discipline with respect to designing a simulation and learning environment that aims to improve teachers' diagnostic competences is of course teacher education. It is important to understand that diagnosing students' psychological problems is only one among many demands teachers are asked to fulfill in their everyday practice. Therefore, realistic learning objectives must first be determined. It seems reasonable to suggest that teachers should be able to identify students' psychological problems in terms of distinguishing between clinically relevant and nonrelevant behavior, reflect on potential hypotheses and generate, evaluate, and integrate evidence obtainable in the everyday school setting. Therefore, the learning goal is the capability to draw substantiated conclusions and formulate argumentation texts to communicate these conclusions to other teachers and psychological professionals.

The distinction between clinically relevant and nonrelevant behavior and the classification of symptoms in terms of disorders are closely related to the discipline of clinical psychology. These concepts build on diagnostic categories defined by the medical domain and documented in diagnostic manuals such as the ICD-10 (Dilling et al., 2015), which serves as the diagnostic reference standard in Germany. To achieve the aforementioned learning goal, pre-service teacher education needs to provide basic conceptual knowledge on diagnostic classifications and related

symptomatology that are particularly relevant for the age group served by a given school type. Moreover, some general strategic knowledge on how to approach diagnosing, generate evidence, and differentiate between different diagnoses with similar manifestations is necessary.

To design an effective learning environment that targets teachers' diagnostic competences, research on diagnostic processes and actions should be taken into account. Such research can primarily be found in the discipline of medical education. A central insight in this field is that learning how to apply conceptual diagnostic knowledge and diagnostic strategy based on case information requires repeated practice (Schmidt & Rikers, 2007). In medical education, this practice is commonly provided by confronting learners with virtual patients (Berman et al., 2016). Educators present virtual patients in different presentation formats. One such format is the serial cue format, which presents case information separated by units. Typically, the case information is presented as the results of various medical tests, which can be accessed in a sequential fashion.

8.4 Simulation Description

FAMULUS designs and tests a learning environment involving document-based simulation to foster diagnostic competences. The learning environment is implemented using the learning management system CASUS (Simonsohn & Fischer, 2004). Building on the idea of virtual patients, the learning environment presents six cases of students showing problems that are potentially related to a behavioral, developmental or learning disorder. The cases were developed with the involvement of experts in school psychology and educational sciences. Blueprints were created before the case information was divided up and assigned to informational sources like "classroom observation" or "meeting with parents". Based on the blueprints, different types of learning materials were developed, e.g., written records of conversations or observations and visuals of documents, such as report cards and school assignments. Following this procedure, six cases in the serial cue format were designed and implemented in the simulation-based learning environment. Another expert from psychotherapy validated the cases in terms of symptomatic authenticity and representativeness.

During the learning phase, learners first watch a 20-min video presenting basic knowledge about diagnosing and behavioral, developmental, and learning disorders among students. This video was included to meet learners' prerequisites (see Chernikova et al., 2022) by addressing their limited prior professional knowledge base. Next, learners are asked to adopt the perspective of a teacher and diagnose the six learning cases. While interacting with the learning environment, they need to apply four epistemic activities in particular (Chernikova et al., 2022; Fischer et al., 2014): generating hypotheses, generating evidence, evaluating evidence and drawing conclusions. For each case, they receive brief initial problem information. On this basis, the learners need to generate up to three initial hypotheses. They then can access the complete case information, which is presented in

serial cue format with the following informational sources: the teacher's classroom observations, schoolyard observations, school assignments and report cards as well as conversations with other fictional teachers, the student him- or herself and the student's parents. The learners do not have to examine all informational sources but make selections and stop the information search at any time. Thus, the learning environment simulates the activities of evidence generation and evaluation. As a final task for each case, learners have to draw a diagnostic conclusion. Moreover, they are asked to communicate their diagnostic actions and write a substantiated argumentation text for their conclusion in a free-text format.

8.5 Feedback Description

As part of the learning environment, an automatic adaptive feedback tool was designed as a learner support (see Chernikova et al., 2022). It specifically addresses the gap between a learner's answer and the sample solution for each learning case and provides hints on how to better apply relevant conceptual and strategic knowledge. Providing such process-related explanations which point the learners to individual options for improvement has been shown to be more effective for learning competences than simpler feedback like presenting the correct response—e.g., an expert solution (Hattie & Timperley, 2007).

Learners receive automatic adaptive feedback on their diagnostic argumentation texts. The feedback is given on two levels: the application of a diagnostic strategy and the application of case-specific concepts. The general diagnostic strategy refers to the epistemic activities of generating hypotheses, generating evidence, evaluating evidence and drawing conclusions (Fischer et al., 2014). The case-specific concepts concern differential hypotheses in the clinical spectrum (e.g., ADHD) as well as hypotheses in the nonclinical spectrum (e.g., family problems), and particular evidence (e.g., inattention, hyperactivity and impulsivity). To provide in-time automatic adaptive feedback, the learners' argumentation texts are automatically analyzed by an NLP algorithm, more specifically an artificial neural network (Schulz et al., 2019). The algorithm automatically identifies the presence or absence of the four epistemic activities and several case-specific concepts. It does so by calculating the likelihood of expressions' affiliation to previously trained categories. This enables the algorithm to automatically analyze new texts and recognize unknown expressions, which, however, need to be similar to what the algorithm learned earlier. This automatic analysis, in turn, activates a range of predefined feedback components. These components combine to form a real-time automatic adaptive feedback response for each learner's argumentation text for each learning case.

If, for example, a learner did not draw a diagnostic conclusion in their argumentation text, he or she receives the feedback that this component is essential but missing in their submitted argumentation text. The learner is also prompted to include a substantiated conclusion in their next argumentation text. One example for feedback on the conceptual level would be the confirmation of correctly considered diagnoses and the correction of incorrectly considered diagnoses as well as feedback on specific evidence used to justify the arguments.

The overall quality of the adaptive feedback critically depends on how accurately the NLP algorithm detects epistemic activities and case-specific concepts. The following section further illustrates the associated tasks and challenges for the project, referring to the example of automatically analyzing epistemic activities.

8.6 Training of an NLP Algorithm

Previous studies have already attempted to train NLP algorithms for the automatic identification of epistemic activities (Daxenberger et al., 2018). These algorithms were trained based on the coding of think-aloud protocols of pre-service teachers diagnosing everyday classroom problems (Csanadi et al., 2016) and social workers diagnosing client problems (Ghanem et al., 2018). These studies applied the method of conditional random fields (CRFs; Okazaki, 2007). CRFs consider the correlations between adjacent codes to identify the best chain of codes for each sentence (Ma & Hovy, 2016). However, the accuracy of the algorithms in identifying epistemic activities was rather weak.

The FAMULUS algorithm is trained based on argumentation text data collected in the context of a previous study. This previous study had 118 pre-service teachers learn with a preliminary version of the FAMULUS simulation-based learning environment involving the six current learning cases and two additional cases from the same symptomatic spectrum. The resulting data set of 944 argumentation texts was manually coded by four coders concerning the four epistemic activities of generating hypotheses, generating evidence, evaluating evidence and drawing conclusions. The intercoder reliability was calculated based on 150 fourfold-coded texts, resulting in sufficient agreement.

Based on a data set of 440 argumentation texts, a first neural network model was fitted. The CRF method was combined with the more recent method of bidirectional long short-term memory (BiLSTM; Reimers & Gurevych, 2017). The BiLSTM method considers the overall context of codes within the text by looking at bidirectional long-term dependencies (Ma & Hovy, 2016). Schulz et al. (2019) provide further details about the methodology and model fitting process.

The performance of the algorithm was tested on 110 additional argumentation texts, showing a satisfactory model fit. The algorithm's coding performance was also compared to the human intercoder reliability and achieved more than 70% of the human coding performance. Moreover, the FAMULUS algorithm achieved almost twice the performance reported by previous studies attempting to train algorithms for the automatic identification of epistemic activities (Daxenberger et al., 2018).

In the future, the training data set for the algorithm will be extended to the full data set of 944 argumentation texts. The algorithm will also be extended to automatically code the dimension of case-specific concepts. The extended and improved algorithm will then serve as a basis for the automatic adaptive feedback component.

8.7 Outlook

In an upcoming laboratory study, the automatic adaptive feedback will be compared with a nonadaptive feedback option regarding the effect on learning diagnostic competences. In doing so, we will contribute to Questions 2 (learner support) and 4 (adaptation) of the overarching research questions mentioned in the introduction by Fischer et al. (2022). The proposed sample for the study consists of 180 pre-service teachers. They will access the learning environment, diagnose the six simulated learning cases and write an argumentation text for every case. Participants in the experimental condition will receive adaptive feedback in line with their argumentation texts, while participants in the control group will receive static feedback consisting of a comprehensive expert solution. The effects of both types of feedback will be analyzed regarding several outcomes: (1) diagnostic accuracy in the learning cases and (2) knowledge gain from pre- to post-test. It is expected that the automatic adaptive feedback will exceed the nonadaptive expert solution in terms of participants' performance and learning gains.

This experimental study will be replicated in a second FAMULUS sub-project that develops a highly similar learning environment and adaptive feedback component to foster diagnostic competences in medical education. In the medical learning environment, learners will have to diagnose six patients with symptoms of fever or back pain. An interdisciplinary comparison of the sub-projects from teacher and medical education regarding learners' interactions with the learning environment and the structure of their diagnostic argumentations might reveal interesting results as well. One example would be to explore sequences of epistemic activities in diagnostic argumentation (see Csanadi et al., 2018). The sequence of epistemic activities seems to differ substantially across pre-service teachers. A comparison with medical students might indicate interdisciplinary similarities or differences in the variability and predominant patterns of sequences. Moreover, changes in variability and sequences across the learning cases will be examined.

Another area of exploration is how to further improve the NLP algorithm's coding accuracy. The accuracy generally depends on several determinants, such as the consistency and quality of the text material, the amount of training data, the consistency and quality of the training data, and consistency of the coding in the training data. One solution approach within the FAMULUS project is an attempt to improve the consistency and quality of the text material that has to be coded. The previously collected text material currently being used as training data will be analyzed in terms of the potential need to further clarify the task instructions. Improving the instructions (if necessary) might in turn improve the consistency and quality of the argumentation texts collected in the upcoming study and hence future additional NLP training data. Adding argumentation texts from the upcoming study to the training data will also increase the overall amount of training data. These steps might further increase the algorithm's accuracy and thus also the quality and effectiveness of the automatic adaptive feedback.

Lastly, it will be interesting to examine how the FAMULUS learning environment can be integrated into actual higher education classes. This transfer will be

investigated in a field study. This simulation-based learning opportunity will be offered in regular teacher education classes at three different universities. The implementation will be evaluated and the results of the laboratory studies will be validated.

8.8 Conclusion

Simulation-based learning is a feasible approach to implement effective learning environments in higher education for competences, such as diagnostic competences. However, learning competences requires specific and intensive learner support. Implementing high-quality learner support that can be feasibly applied on a large scale is a major challenge. Automation using artificial intelligence seems to be a promising way to approach some parts of these challenges. FAMULUS illustrates and evaluates natural language processing measures to automate process-related feedback on diagnostic argumentation text answers. Some initial applications of the natural language processing algorithms presented in this chapter indicate that the automated text analyses might be sufficiently accurate to support learners with adaptive process-related feedback during their learning. This appears to be particularly important in interdisciplinary and ill-defined fields of application like teachers' diagnosing of students' behavioral, developmental, and learning disorders, where corresponding learning opportunities are largely lacking or neglect competence-oriented learning.

Acknowledgments This research was supported by a grant from the German Federal Ministry of Research and Education (Grant No.: 16DHL1039).

References

- Ananiadou, K., & Claro, M. (2009). *21st century skills and competences for new millennium learners in OECD countries*. OECD.
- Baek, Y. (2009). Digital simulation in teaching and learning. In *Digital simulations for improving education: Learning through artificial teaching environments* (pp. 25–51). IGI Global.
- Berman, N. B., Durning, S. J., Fischer, M. R., Huwendiek, S., & Triola, M. M. (2016). The role for virtual patients in the future of medical education. *Academic Medicine*, *91*(9), 1217–1222.
- Chernikova, O., Heitzmann, N., Opitz, A., & Fischer, F. (2022). A theoretical framework for fostering diagnostic competences with simulations. In F. Fischer & A. Opitz (Eds.), *Learning to diagnose with simulations—Examples from teacher education and medical education*. Springer.
- Coderre, S., Mandin, H., Harasym, P. H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education*, *37*(8), 695–703.
- Codreanu, E., Huber, S., Reinhold, S., Sommerhoff, D., Neuhaus, B., Schmidmaier, R., & Seidel, T. (2022). Diagnosing mathematical argumentation skills: A video-based simulation for pre-service teachers. In F. Fischer & A. Opitz (Eds.), *Learning to diagnose with simulations—examples from teacher education and medical education*. Springer.

- Csanadi, A., Kollar, I., & Fischer, F. (2016). *Scientific reasoning and problem-solving in a practical domain: Are two heads better than one?* International Society of the Learning Sciences.
- Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 13(4), 419–438.
- Daxenberger, J., Csanadi, A., Ghanem, C., Kollar, I., & Gurevych, I. (2018). Domain-specific aspects of scientific reasoning and argumentation: Insights from automatic coding. In *Scientific reasoning and argumentation* (pp. 44–65). Routledge.
- Dilling, H., Mombour, W., & Schmidt, M. H. (2015). Internationale Klassifikation psychischer Störungen. In *ICD-10 Kapitel V (F). Klinisch-diagnostische Leitlinien* (10th ed.). Hogrefe.
- Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review*, 22(1), 89–102.
- Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., & Earhart, J. (2009). Early identification of behavioral and emotional problems in youth: Universal screening versus teacher-referral identification. *The California School Psychologist*, 14, 89–95.
- Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Fischer, M., et al. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, 2(3), 28–45.
- Fischer, F., Chernikova, O., & Opitz, A. (2022). Learning to diagnose with simulations: Introduction. In F. Fischer & A. Opitz (Eds.), *Learning to diagnose with simulations—examples from teacher education and medical education*. Springer.
- Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., Schmidmaier, R., et al. (2018). Systematizing professional knowledge of medical doctors and teachers: Development of an interdisciplinary framework in the context of diagnostic competences. *Educational Sciences*, 8(4), 207.
- Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology*, 45(6), 1097–1114.
- Ghanem, C., Kollar, I., Fischer, F., Lawson, T. R., & Pankofer, S. (2018). How do social work novices and experts solve professional problems? A micro-analysis of epistemic activities and the use of evidence. *European Journal of Social Work*, 21(1), 3–19.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hölling, H., Schlack, R., Petermann, F., Ravens-Sieberer, U., Mauz, E., & Group, K. S. (2014). Psychische Auffälligkeiten und psychosoziale Beeinträchtigungen bei Kindern und Jugendlichen im Alter von 3 bis 17 Jahren in Deutschland—Prävalenz und zeitliche Trends zu 2 Erhebungszeitpunkten (2003–2006 und 2009–2012) [Psychological and psychosocial problems of German children and adolescents from age of 3 to 17—Prevalences and temporal trends of 2 measurement points (2003–2006 und 2009–2012)]. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 57(7), 807–819.
- Lawson, A. E., & Daniel, E. S. (2011). Inferences of clinical diagnostic reasoning and diagnostic error. *Journal of Biomedical Informatics*, 44(3), 402–412.
- Ma, X., & Hovy, E. (2016). End-to-end sequence labeling via bi-directional lstm-cnns-crf. arXiv preprint arXiv:1603.01354.
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499–515.
- Naser, S. S. A. (2012). Predicting learners performance using artificial neural networks in linear programming intelligent tutoring system. *International Journal of Artificial Intelligence & Applications*, 3(2), 65.
- Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, 35(5), 501–517.
- Okazaki, N. (2007). *CRFsuite: A fast implementation of conditional random fields*. Retrieved from <http://www.Chokkan.Org/software/crfsuite>
- Papandrea, K., & Winefield, H. (2011). It's not just the squeaky wheels that need the oil: Examining teachers' views on the disparity between referral rates for students with internalizing versus

- externalizing problems. *School Mental Health*, 3(4), 222–235. <https://doi.org/10.1007/s12310-011-9063-8>
- Reimers, N., & Gurevych, I. (2017). Reporting score distributions makes a difference: Performance study of lstm-networks for sequence tagging. arXiv preprint arXiv:1707.09861.
- Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, 26(1), 1.
- Roth, D. M., Leavey, G., & Best, R. (2008). On the front-line: Teachers as active observers of pupils' mental health. *Teaching and Teacher Education*, 24(5), 1217–1231.
- Schmidt, H. G., & Rikers, R. M. (2007). How expertise develops in medicine: Knowledge encapsulation and illness script formation. *Medical Education*, 41(12), 1133–1139.
- Schrader, F.-W. (2011). Lehrer als Diagnostiker [teachers as diagnosticians]. In *Handbuch der Forschung zum Lehrerberuf* (pp. 683–698).
- Schulte-Körne, G. (2016). Mental health problems in a school setting in children and adolescents. *Deutsches Ärzteblatt International*, 113(11), 183.
- Schulz, C., Meyer, C. M., & Gurevych, I. (2019). Challenges in the Automatic Analysis of Students' Diagnostic Reasoning. In A. Korhonen, D. Traum, & L. Márquez (Eds.), *The 57th Annual Meeting of the Association for Computational Linguistics - proceedings of the conference: July 28-August 2, 2019, Florence, Italy* (pp. 6974–6981). Association for Computational Linguistics.
- Simonsohn, A. B., & Fischer, M. R. (2004). Evaluation of a case-based computerized learning program (CASUS) for medical students during their clinical years. *Deutsche Medizinische Wochenschrift*, 129(11), 552–556.
- Spinath, B. (2005). Akkurtheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz [accuracy of teacher judgments on student characteristics and the construct of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie*, 19(1/2), 85–95.
- Trudgen, M., & Lawn, S. (2011). What is the threshold of teachers' recognition and report of concerns about anxiety and depression in students? An exploratory study with teachers of adolescents in regional Australia. *Journal of Psychologists and Counsellors in Schools*, 21(2), 126–141.
- Van Merriënboer, J. J., Clark, R. E., & De Croock, M. B. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology Research and Development*, 50(2), 39–61.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

