PERFORMANCE ASSESSMENT AND DIGITAL TRAINING FRAMEWORK FOR YOUNG PROFESSIONALS' GENERIC AND DOMAIN-SPECIFIC ONLINE REASONING IN LAW, MEDICINE, AND TEACHER PRACTICE

EVALUACIÓN DEL RENDIMIENTO Y ENTRENAMIENTO DIGITAL PARA EL RAZONAMIENTO CRÍTICO GENÉRICO Y ESPECÍFICO PARA JÓVENES PROFESIONALES EN LA FASE PRÁCTICA EN EL DERECHO, LA MEDICINA Y LA DOCENCIA

Olga Zlatkin-Troitschanskaia
Sebastian Brückner
Marie-Theres Nagel
Ann-Kathrin Bültmann
Jennifer Fischer
Susanne Schmidt
Dimitar Molerov

ABSTRACT

In the digital age, the Internet is increasingly considered a major information source. This is especially true for informal, e.g., post-university, learning. Evidentially, young professionals are increasingly using online sources as an information and learning tool. Critical reasoning from online information for learning and professional processes in the domains of medicine, law, and teaching is considered a highly relevant competence facet. For example, staying up to date on a multitude of matters, e.g., published in articles and guidelines, as is the case in the medical field, can be challenging when the required competencies to use online media are absent (e.g., Allen et al. 2005, O'Carroll et al. 2015). Current research on students in higher education indicates substantial deficits in their critical online reasoning skills, also among graduates. However, online information seeking and corresponding competencies of young professionals in job-specific educational processes have not been researched yet. There is a lack of both valid domain-specific assessments for different professions and learning tools that can effectively foster the competent use of online information in practice among young professionals.

Our research presented here is part of the collaborative BRIDGE project, which is conducted under the umbrella of the program "Research for the Design of Educational Processes under the Conditions of Digital Change." This study is based on our previous work on the assessment of generic skills in higher education in the international projects CLA+, iPAL, and CORA as well as on experiences with job-specific performance assessments from the research programs KoKoHs and ASCOT+, which measured professional competence. To validly measure critical online reasoning among young professionals from three domains — medicine, law, and teacher training —we develop new computer-based online performance assessments and corresponding training tools. The specific aim is to analyze to what extent they improve in using online information with greater reflection when creating job-specific documents after an online training based on process and performance data (using innovative approaches, such as text mining and educational data mining). In this paper, we showcase the conceptual and assessment framework of the newly developed innovative tools to measure and promote generic and domain-specific online reasoning among young professionals in medicine, law, and teacher education. Based on this framework, we discuss how these crucial professional competence facets can be validly measured and effectively fostered in practice.

Key words: Digital competencies, critical online reasoning, job-related education, digital training tools, performance assessments

RESUMEN

En la era digital, Internet se considera cada vez más como una importante fuente de información. Esto es especialmente cierto para el aprendizaje informal, como por ejemplo, la educación postuniversitaria. Evidentemente, los jóvenes profesionales utilizan cada vez más las fuentes en línea como herramienta de información y aprendizaje. El razonamiento crítico de información en línea para el aprendizaje y los procesos profesionales en la medicina, el derecho y la docencia se considera una faceta de competencia de gran relevancia. Por ejemplo, mantenerse actualizado sobre una multitud de asuntos tales como publicados en artículos y pautas, como es el caso en la medicina, puede ser un desafío cuando no están desarrolladas las competencias necesarias para utilizar los medios en línea (Allen et al. 2005, O'Carroll et al.2015). Las investigaciones actuales sobre estudiantes universitarios indican deficiencias sustanciales en sus habilidades de razonamiento crítico en línea, también entre los graduados. Sin embargo, todavía no se han investigado la búsqueda de información en línea y las competencias correspondientes entre los jóvenes profesionales en la fase práctica. Hay una falta tanto de evaluaciones válidas específicas para diferentes profesiones como de herramientas de entrenamiento que puedan fomentar de manera efectiva el uso competente de la información en línea entre los jóvenes profesionales en la práctica. Nuestra investigación es parte del proyecto colaborativo BRIDGE, que hace parte del programa "Investigación para el Diseño de Procesos Educativos en las Condiciones del Cambio Digital". Este estudio se basa en nuestro trabajo anterior sobre la examinación de habilidades genéricas en la educación superior en los proyectos internacionales CLA+, iPAL y CORA, así como en experiencias con exámenes de rendimiento específicas para profesiones acumuladas en los programas de investigación KoKoHs y ASCOT+, que evaluaron competencias profesionales. Para examinar válidamente el razonamiento crítico en línea entre los jóvenes profesionales (en el derecho, la medicina y la docencia), desarrollamos nuevos exámenes de rendimiento en línea y entrenamientos digitales correspondientes. El objetivo es analizar en qué medida los jóvenes profesionales mejoran utilizando la información en línea con mayor reflexión mientras su preparación de documentos profesionales después de haber participado en un entrenamiento en línea. Utilizamos datos del proceso y del rendimiento (utilizando enfoques innovadores, como la minería de textos y la minería de datos educativos). En este artículo, mostramos el marco conceptual y de evaluación de los instrumentos innovadores recientemente desarrollados para medir y fomentar el razonamiento crítico en línea genérico y especializado entre los jóvenes profesionales en la fase práctica de la medicina, del derecho y de la docencia. Basado en este marco, discutimos cómo estas facetas de competencia profesional importantes se pueden medir de manera válida y fomentar de manera efectiva en la práctica.

Key words: Competencias digitales, razonamiento crítico, educación postuniversitaria, herramientas de entrenamiento en línea, evaluación de rendimiento.

Date of receipt: 15 april 2021
Date of acceptance: 9 june 2021

1. INTRODUCTION

1.1. RESEARCH BACKGROUND AND FOCUS

In the 21st century, the increasing digitalization of education has been shaping the learning and professional processes and has led to the development of a new 'cultural-technical skill': a competent use of online information (Redecker, 2017). This skill plays an increasingly central role in the required skill set in several professions. This is particularly evident in the professional education of prospective medical, teaching, and legal trainees (Kuhn et al., 2020), professions for which evidence-based decision-making and acting is especially required in practice.

Several studies have shown that students and prospective young professionals not only use digital media in their everyday lives but increasingly use the Internet or social media for learning (Brooks, 2016; Kimmerle et al., 2015; Maurer et al., 2020). Increases in Internet use, however, are not automatically accompanied by increases in competency in dealing with digital information as many current studies indicate (McGrew et al., 2019; Wineburg et al., 2018). Undergraduate and graduate students often lack the skills to critically examine and evaluate the information they find online (McGrew et al., 2017), although these represent a central practice-relevant competence facet, especially in the professional domains of medicine, law, and teaching, as these professional groups in particular have a significant social impact and increasingly use online research for work purposes (for examples, see Section 2.1).

Due to these competence deficits, the question arises whether and to what extent the ability to critically use online information can be fostered in professional education and practice. Recent studies have indicated that this kind of ability can be trained in higher education (McGrew et al., 2019; Weber et al., 2018; 2019). However, the range of training approaches and opportunities available to date has been severely limited. Existing tools, e.g., are limited to a mere list of questions about the website visited for information (McGrew et al., 2017). Working through the items on a checklist, however, does not provide the necessary critical analysis but only leads to a superficial examination (Wineburg & McGrew, 2018). Hence, approaches of this kind fall short as they no longer do justice to the complexity of the Internet and social media (Wineburg et al., 2018). Overall, there is a lack of learning tools that can effectively foster the competent use of online information in practice among young professionals, as well as a lack of valid domain-specific assessments of this skill for different professions.

The research from the BRIDGE project presented here addresses this significant gap by developing and implementing a holistic competence framework and technology-based approach to examine and support the critical use of online media among professionals in training contexts.

Our research in the collaborative BRIDGE project, funded by the German Federal Ministry of Education and Research (BMBF) and conducted under the umbrella of the newly established federal program "Research for the Design of Educational Processes under the Conditions of Digital Change,"[1] is based on our previous work on the assessment of generic skills in higher education in the international projects CLA+ (Zlatkin-Troitschanskaia et al., 2018), iPAL (Shavelson et al., 2019), and CORA (Molerov et al., 2020) as well as on experiences with job-specific performance assessments from the research programs KoKoHs (Zlatkin-Troitschanskaia et al., 2020) and ASCOT+ (BMBF, 2015), which measured professional competence.

[1] This new Federal research program focuses "on the design of educational processes under the conditions of digital change" and therefore lies on both research and development of new digital educational technologies, which includes: (1) developing new innovative concepts for the design of educational processes that use the potential of digital media to support individual and collective educational success, (2) generating scientifically sound knowledge for the effective use of digital concepts and tools in professional practice, and (3) tapping the future potential of digital media in individual and collective educational processes.

The interdisciplinary collaborative project BRIDGE was launched in the summer of 2020 and will extend for three years. As a part of this research program, we aim to comprehensively examine the development of the critical online reasoning skills among young professionals in the domains of medicine, law, and teacher education in economics that are required to effectively use online information in professional settings and how these skills can be fostered using a digital training. To this end, we develop and implement both an innovative online performance assessment of these skills and a corresponding digital training concept in the post-academic professional training phase in these three domains.

One particular research focus lies on controlling for indicators and influencing factors relevant to study participants' critical online reasoning skills as well as to the effectiveness of new digital tools, such as participants' personal and contextual (e.g., job-related) characteristics. Based on a longitudinal assessment, changes in their professional use of digital media in the practical training phase are analyzed in a multi-step approach considering both performance data (e.g., assessment results) and process data (e.g., log files with time stamps). After every survey, the results are reported back to the study participants and cooperation partners in educational practice through systematic feedback. Finally, the identified potential of the newly developed digital tools for the promotion of professional use of digital media will be distributed through collaboration with partners in research and educational practice.[2]

In this paper, we showcase the conceptual and assessment framework of the newly developed innovative computer-based online performance assessments to measure generic and domain-specific online reasoning among young professionals in medicine, law, and teacher education and the corresponding training tools to enhance the critical use and reflection of online information during job-related activities of young professionals (e.g., creating job-specific documents) from these three domains. Based on this framework, we discuss how these crucial professional competence facets can be validly measured and effectively fostered in practice.

1.2. RESEARCH OBJECTIVES

Building on prior research in this field (see Section 2.1) as well as on our preliminary studies (see Section 2.2), the objectives in the BRIDGE project include:

(1) Analyzing and promoting the general critical use of online information (hereafter: generic critical online reasoning (GEN-COR; for details, see Section 2.2) among first-year prospective law clerks, economics teachers and medical interns/doctors in an ecologically valid longitudinal design.

(2) Implementing a newly developed digital training approach/tool to promote GEN-COR through online training (as an informal learning opportunity) and analyzing its effectiveness in a pre-post design.

(3) Analyzing the domain-specific use of online information when creating typical job-specific documents in everyday professional practice (lesson/teaching plans, clinical or legal case reports etc.) (hereafter: domain-specific critical online reasoning (DOM-COR; for details, see Section 2.2); DOM-COR is directly and validly captured both via process data (such as event log data, browsing history, eye movements) during document creation and via performance data, i.e., the (semi-automated) scoring of the created documents. The data analysis is qualitatively enriched through a subject-specific evaluation by partners from educational practice (e.g., instructors).

(4) Implementing a newly developed digital assessment and analyzing personal (e.g., prior education, intelligence) and contextual factors (e.g., the type/number of courses completed, e.g., at a teacher training college) that may influence the development of GEN- and DOM-COR and analyzing how/to what extent they can be fostered.

[2] This project is being conducted in cooperation with researchers at the University of Frankfurt as well as numerous cooperation partners. For more information, see the project's homepage https://eng.bridge.uni-mainz.de/.

(5) Analyzing the interplay between the informal promotion of GEN-COR through a newly developed digital training and the formal learning opportunities completed during the practical phase, as well as their effect on the development of GEN- and DOM-COR.
(6) Analyzing specific influences of professional practice on DOM-COR in a multitrait-multimethod design (MTMM; Campbell & Fiske, 1959) by examining the three groups (prospective judges, doctors, and economics teachers) comparatively, as well as in terms of their GEN-COR skills using the method of known groups (Hubley & Zumbo, 2011).
(7) Developing evidence-based (media-)didactic recommendations for the professional practice phase to promote DOM-COR in the three domains as well as GEN-COR across domains.
Overall, the development and promotion of GEN-COR and DOM-COR within and outside of institutional learning settings in professional practice are explored in a domain-specific and comparative manner among the three domains, and promising interactions between formal (offline) and informal digital learning opportunities are examined based on the conceptual, assessment, and training frameworks presented below.

2. THEORETICAL AND CONCEPTUAL BACKGROUND FOR ASSESSING AND PROMOTING COR

2.1. CURRENT STATE OF RESEARCH

In the context of increasingly location-/time-independent, individual, multimodal, and self-directed learning processes in the digital environments, the competent and effective use of online media and information plays a central role (National Research Council, 2012; Wiley et al., 2009). These skills are also regarded as a crucial facet of professional competence in the fields of medicine, law, and teaching. For instance, several studies show an increasing significance of the Internet and COR skills in medical professional practice (e.g., Mesko & Győrffy, 2019). For example, a study on media literacy in medical studies and preparation for the later professional field in medicine (Rott, 2014) shows that "the Internet is an important tool for research, in which an overview article with the most important information can be found about every known disease". In addition to conducting research in professional databases, there are also clinical pictures for which "a search with Google finally provides decisive insights into the cause of the disease. The expert use of databases and search engines is therefore essential in this professional field" (Roth 2014, p. 160; see also Watson & Burr, 2018; Kuhn et al., 2020).
Law presents a similar case. For instance, law students repeatedly list online sources of questionable credibility in their exams (Basak & Schimmel, 2008). When assessing printed literature, law students know which sources to use, "but when evaluating Internet sources, this feeling is still surprisingly lacking in many students" (ibid., p. 436). Thus, the problem seems to lie less in poor legal research skills, and rather in the selection of unprofessional sources and lacking cross-checking of information found on the Internet, as students in law school are not trained to search for evidence-based online information (Schimmel, 2011). At the same time, online media are now more popular than traditional media for researching information, even among law students (Mielke & Wolff, 2007; see also Meredith, 2010; Wagner, 2018).
As professional knowledge brokers, teachers need COR skills not only to appropriately use online information in their everyday work at school, but are also expected to convey skills of competent information use to students in their classrooms (Schiefner-Rohs, 2012). Hence, the ability to deal critically with online media is of particular significance in this profession. At the same time, several studies show that teachers lack "the technical know-how and the necessary methodological-didactic Internet competence" (Maireder & Nagl, 2010, p. 5) as well as a critical-reflexive approach in their handling of (online) sources and information (see also Amin, 2016; Hague & Payton, 2010).

In addition to their personal need for competent use of online information, professionals in all three domains also receive lay people (clients, patients, students) who inform themselves about medical or legal issues, or learning-related topics online. Knowing how to competently respond to clients, patients, and students is part of successful communication in the respective profession, which is complicated by diverse knowledge lay people have access to online (medical communication surrounding diverse (mis)information during the pandemic is only a recent example). Hence, in a mid-term perspective, after mastery of COR skills, transfer to understanding COR from a lay perspective becomes key for professionals as well.

From the productive side, part of professional communication itself has moved online, e.g., in the form of web doc wikis or lawyer's blogs on specific legal questions. Knowing about cues of (un)trustworthy online information in general and in their field can help professionals more successfully communicate their services online.

Despite this high degree of practical relevance, only little research has been conducted on the use of online media and information, the development of corresponding skills throughout education, and how they can be effectively promoted in professional training processes in post-university education (Steffens et al., 2017; for an overview, see Zlatkin-Troitschanskaia et al., 2021). Approaches and findings from existing studies on other learner groups, such as schoolchildren and higher education students, cannot be directly transferred to young professionals due to different job-specific requirements and situations with regard to digital media use.

Moreover, existing studies have various conceptual and methodological limitations. As a current literature review of over 500 international studies indicates (Zlatkin-Troitschanskaia et al., 2021), most studies on students' media use are still based on self-assessments and one measurement point and, therefore, do not allow for any insights into the development of such skills. They often only focus on declarative knowledge facets or beliefs, but not on actual competencies and students' performance in dealing with online media. Only few studies capture online information use in a realistic online environment (in real time) while gathering both process and performance data and assessing development over time using pre-post-measurements. Only very few studies applied intervention designs and investigated how online media use can be promoted in educational practice (Booth, 2011; Mathson & Lorenzen, 2008; for process-related studies, Tossell et al., 2015; Gadiraju et al., 2018; Yu et al., 2018; Goldhammer & Zehner, 2017; for instruments for assessing digital media competencies and for relevant concepts, such as digital literacy, see, e.g., Reichert et al., 2020). With regard to the promotion of competent media use, only few intervention studies were found; none focused on young professionals or practical training phases. To promote competent use of generic online information, Weber et al. (2019) and McGrew et al. (2019) developed training tools and demonstrated their effectiveness in intervention studies with higher education students in on-site classes. However, no studies were found on the effectiveness of digital training courses in post-university job-related contexts.

The studies conducted with students at the end of their bachelor's/master's programs (incl. medicine and economics education, which end with the first state examination) as well as with university graduates point to significant deficits. These deficits became apparent both in students' general use of sources of information (McGrew et al., 2018; Nagel et al., 2020a; Schmidt et al., 2020) as well as domain-specific competence facets (e.g., question-guided Internet searches, selection and evaluation of websites, avoidance of unreliable sources, selection of information, incorporation of information into one's argumentation, synthesis of evidence/information in response to a question, communication of results, and the provision of evidence) during their preparation of practice-relevant professional documents (Zlatkin-Troitschanskaia et al., 2019; Nagel et al., 2020b).

Overall, no studies on the use of online information among young professionals in law and among medical interns were found. However, the preliminary research (Nagel et al., 2020a; Molerov et al.,

2020; Schmidt et al., 2020; Banerjee et al., 2020) indicates that COR is of great importance for successful learning in digital contexts and that its facets (as described in Section 2.1), including the ability to critically evaluate online information, can be developed and trained in a targeted manner.

2.2. CONCEPTUAL FRAMEWORK OF COR

Our research in BRIDGE addresses these deficits by focusing on the development of work-related use of online information in the context of professional training by capturing and analyzing media use in a direct, valid, and process-/performance-based manner and analyzing how it can be promoted using a newly developed digital tool.

To build a sound conceptual basis, we adapted and further developed the construct *'civic online reasoning',* developed by McGrew et al. (2017, 2019). To measure and to promote this construct, which focuses on the critical evaluation of online information, McGrew et al. (2017) developed a new assessment tool and a corresponding instructional intervention (McGrew et al., 2019). The assessment focuses particularly on the ability to evaluate digital content and make a reasoned decision regarding its trustworthiness.

McGrew et al. (2019, p. 486) use the term 'civic online reasoning' to describe the ability to effectively seek, evaluate, and verify social and political information provided online. Guiding questions for successful performance included: "Who is behind the information?", "What evidence is there?", and "What do other sources say?" (McGrew et al., 2017, p. 5). These questions coincided with the approach professional fact-checkers take in checking the reliability of sources of information. For instance, to answer the question, "Who is behind the information?", fact-checkers use a reading technique called 'lateral reading' (McGrew et al., 2019, p. 486). While common readers usually tend to read the content offered on web pages in great detail straight away ('close reading'), professional fact-checkers usually leave the web page immediately to research the entities behind the website in a new tab. Only after they have found more detailed information about the web page under investigation and its content as well as perspectives from other sources, do they return to the original web page. Here, a phenomenon becomes evident that can be subsumed under the term 'reliability paradox': "To learn more about a website, you have to leave it" (McGrew et al., 2017, p. 8).

To allow this American assessment to be used in a German context as well, we adapted it for tertiary education in Germany (for details, see Molerov et al., 2020). Due to (socio)cultural embeddedness of US-American items and the use of real websites, mere translation was not sufficient, but a functionally equivalent redesign of the assessment proved feasible (though several limitations emerged with regard to creating a functional German equivalent, (see Molerov et al., 2020). During specification of the key construct criteria for redesign, we also further generalized the focus to encompass not only civic issues, but topics related to students' studies in higher education for which they may acquire information online (e.g., from colleagues within the profession), and additionally, further specified the facets (for details, Molerov et al., 2020). Based on expert interviews which particularly stressed the Humboldts' tradition of critical thinking in German higher education (Beck, 2020), as well as on a systematic synthesis of relevant and related multifaceted constructs, such as critical thinking (Facione, 1990; Jahn & Kenner, 2018), digital literacy (Ainley et al., 2016), media literacy (Abreu et al. 2017), information literacy (Taylor & Dalal, 2014), critical argumentation (Walton, 2006a), scientific reasoning and argumentation (Fischer et al., 2018) and information problem-solving (Brand-Gruwel et al., 2017), we developed a detailed conceptual framework titled *critical online reasoning* — a crucial skill in competently using online information (COR; for further discussion of related constructs see, Molerov et al., 2020). We defined COR as the skill required for "searching, selecting, accessing, processing, and using online information to solve a given problem or build knowledge from this online information while

critically distinguishing trustworthy from untrustworthy information and reasoning argumentatively based on trustworthy and relevant information from the online environment" (Molerov et al., 2020). Based on the current state of research, we considered COR to be a central factor of critical and analytical thinking when using online information. According to our definition (Molerov et al., 2020), COR comprises three central facets:

- Online Information Acquisition (OIA)
- Critical Information Evaluation (CIE)
- Reasoning based on Evidence, Argumentation, and Synthesis (REAS)

OIA focuses on searching for and acquiring information available online. This involves using search engines and databases, entering search queries, and accessing specific websites. CIE focuses on selecting sources of information and evaluating both the information and their sources. CIE is also required to distinguish between high- and low-quality sources. REAS focuses on the integrative evaluation of previously collected information and on considering different perspectives in an argumentatively coherent manner. This conceptual framework of COR with its three main facets forms the basis for developing assessment items and scoring rubrics to evaluate students' answers (for details, see Molerov et al., 2020).

With regard to OIA, the aspect of searching for information is examined in detail, as it is a central element in successfully (online) information problem-solving. In our previous study, we demonstrate that students use sources of varying quality during their Internet searches. In particular, we found that there is a relationship between the number of pages participants access during the search process and their performance on the COR tasks (Nagel et al., 2020a). Accordingly, accessing a larger number of sources was a key indicator of successful overall evaluation of a web page's trustworthiness (McGrew et al., 2017, p. 8). Consulting a higher number of sources increases the likelihood of gaining a differentiated perspective and obtaining suitable references to back up one's claims (Nagel et al., 2020a). Thus, the number of websites visited can be used as an important indicator for successful acquisition of relevant online information (Hölscher & Strube, 2000; White et al., 2009). The ability to select reliable online sources is integral to the OIA facet (Molerov et al., 2020).

CIE describes the identification and critical evaluation of information within a web source (Molerov et al., 2020). In particular, the aspects of "neutrality and balance in the inclusion of facts" are important indicators of the quality of a source (McQuail, 1991; Urban & Schweiger, 2013). Furthermore, the characteristics of the medium used or information about the author are considered indicators of reliability. Judging the reliability and the quality of information leads students to spend more time on high-quality sources and ignore unreliable sources. This requires them to judge the credibility of the medium. The ability to recognize cues for trustworthiness of online information is therefore an important component of the CIE facet (Molerov et al., 2020).

The REAS facet focuses on the use of evaluated information in constructing a coherent and sound argument, while considering different perspectives (Molerov et al., 2020). It is crucial that learners do not violate standards of reasoning and argumentation when making their case, e.g., observe basic informal logic when drawing inferences (Walton, 2006b) and back up their claims with information from the websites they evaluated. Providing evidence for claims is an essential requirement for scientific work in the academic context (i.e., in online contexts, e.g., URLs are an important indicator for COR analyses). The classification and weighing of different, possibly conflicting pieces of information, positions, and arguments are also of central importance (Molerov et al., 2020).

Based on the COR's validation (Molerov et al., 2020; Nagel et al., 2020; Schmidt et al., 2020), we further differentiate between (meta)cognitive requirements for COR in two contexts: generic (GEN) and domain-specific (DOM) COR (i.e., GEN-COR and DOM-COR) — mainly based on the topic and degree of domain-specificity of the task scenarios and the online information we use

in test settings. So far, concepts such as analytical reasoning have been postulated as generic skills (Alexander et al., 2016). Some studies, however, point to different profiles and developmental paths between domains (Sá et al., 1999; Toplak & Stanovich, 2002). Conceptually, therefore, we assume that DOM-COR and GEN-COR are initially based on the same overarching COR facets, but the combination of these facets, i.e., specific profiles, may be different depending on the development and application of the skillset in academic and professional domains (DOM-COR) or in domain-general contexts (GEN-COR), which we controlled for in BRIDGE. Hence, DOM-COR task prompts focus on domain-specific problems related to reasoning from online information. DOM-COR is expected to enable young professionals to use specialized online information more competently or identify domain-specific cues in GEN-COR tasks, e.g., a participant with high GEN-COR skills may evaluate an article based on general credibility cues, such as explicit authorship, while a participant with high DOM-COR skills may recognize domain-specific errors in presented information, such as implausible figures.

2.3. HYPOTHESES

In terms of working hypotheses, our validation studies (see Section 2.2.) and a few existing intervention studies (Weber et al., 2019; McGrew et al., 2019) would suggest that the two constructs DOM-COR and GEN-COR with their (sub)facets can be effectively promoted in formal and non-formal educational contexts. Moreover, their development can be related to different personal characteristics such as domain-specific knowledge (assessed via knowledge tests), information and communication technologies (ICT) motivation (Senkbeil, 2018), BIG Five personality traits (Soto & John, 2017), NFC (Need for Cognition) (Cacioppo & Petty, 1982), figural intelligence/visual perception (Liepmann et al., 2007), social media efficacy (Hocevar et al., 2014), and Internet-specific epistemic beliefs (Chiu et al., 2013), Critical Reflection (Frederick, 2005), which are controlled for in BRIDGE in initial surveys (see Section 5).
Based on the prior state of research, several COR-level and development-related assumptions are tested, for instance:

- GEN- and DOM-COR are interrelated but empirically separable, and thus, can each be influenced or promoted differently.
- In all three domains, clusters of entrants with high, medium, and low levels of GEN- and DOM-COR can be identified in the pre-post study.
- Study participants in all three clusters and in all three domains improve their performance in GEN- and DOM-COR from measurement t1 to t2 after participating in the newly developed training, with stronger effects in GEN-GOR; participants with low-level COR at t1 improve more on average.
- At t1 and t2, systematic intraindividual, interindividual, interinstitutional (e.g., training location), and domain-specific (e.g., curricular structure) effects emerge. (Although, due to small samples we cannot expect any significant effects here).
- The newly developed digital training with a feedback system shows stronger effects for participants with fewer completed formal learning opportunities in the professional education phase.

In BRIDGE, the assumptions are further differentiated based on the evidence from the first cohorts assessed in 2020, e.g., regarding expected effects of control variables and, in particular, based on the findings from process and performance data, including data about the online sources and information that young professionals used while solving the GEN-COR and DOM-COR tasks.

2.4. COR ASSESSMENT FRAMEWORK

2.4.1. Generic COR

For the objective, valid assessment of GEN-COR, we used a newly developed online assessment based on a framework initially developed by the Stanford History Education Group and validated across the United States (Wineburg et al., 2018). This large-scale U.S. study examines students' critical use of digital information in a realistic online environment, i.e., using authentic websites, unrestricted online searches as well as civic, consumer, and general knowledge topics. This assessment employs shorter tasks, prompting students to evaluate the trustworthiness of a given source of information on a given topic online. This assessment was then also successfully used as a training tool to significantly promote civic online reasoning among U.S. undergraduates in a controlled intervention study (McGrew et al., 2019).

We conceptually adapted and further developed this framework for the German context (Molerov et al., 2020). The German version, CORA (Critical Online Reasoning Assessment), measures higher education students' skills to search for, select from and critically evaluate online sources and reason from evidence on issues presented on mass and social media and websites. In six short performance tasks, spanning 10 minutes each, students are presented with a description of the context, a task prompt, and a web page to evaluate. They are prompted to perform an open web search, judge online information, and write an open-ended response for each task (for a task example, see Figure 1 in the Supplement). The particular difficulty lies in making a judgment or performing an evaluation in a short amount of time while recognizing hidden interests and biased information representations.

The German version was pre-tested in the 2018 summer term and implemented as a computer-based assessment and successfully tested since the beginning of the 2018/19 winter term. Recently, responses from a sample of more than 200 students from many study domains at German universities supported the validity of CORA according to the criteria of the standards for educational and psychological testing (AERA et al., 2014). We validated the CORA tasks with university students of different domains (incl. prospective economics teachers and doctors; for more information on the assessment, see Molerov et al., 2019; 2020).

In CORA, students evaluate the credibility of websites as sources of information by performing an open web search, judging online information, and writing a short, open-ended response for each task. Some sources include common types of misinformation that need to be identified via critical reading or cross-checking. Test takers do not need specialized domain knowledge to solve the CORA tasks. Examples of task requirements include identifying information mixed with advertisements or evaluating a claim made on social media without a link to an authoritative source. The CORA tasks are characterized by an open-ended information environment, requiring students to perform a live, open web search to identify and determine relevant and credible information, e.g., to verify an 'expert' statement. The difficulty, thus, progresses from following more explicit prompts for COR (e.g., finding information and arguments that contradict a proposition), to recognizing implicit markers/cues of a source's unreliability or biases in information, to narrowing down an open pool of information and evaluating several partly supportive (or conflicting) pro-and-con arguments; ultimately, students need to judge and document the evidence. The distractors in the tasks are elaborate, realistic, and play to well-known biases (e.g., confirmation or authority bias; Paul & Elder, 2005). They are marked by, e.g., the vividness of information representations, amount of information, and level of detail or the use of numerical, statistical, and graphical data (without reference).

In addition to their written responses, students' browser histories are recorded during the web search they conduct to solve CORA tasks. Subsequently, they are surveyed on their sociodemographic and media use behavior to obtain additional information about their search and decision-making processes. The CORO task responses are scored by at least two independent, trained human raters to the rating scheme, which was developed based on the COR construct and takes into account students' recognition of biases, decisions, justifications, and sources (for details on scoring, see Molerov et al., 2020). Particularly justifications are screened for errors in judgment and decision-making.

In validation studies, we conducted comprehensive correlation analyses and t-tests using various variables, such as intelligence, domain-specific knowledge, a scale 'need for cognition,' and information-processing heuristics (which we also controlled for in BRIDGE, see Section 2.3). In terms of cognitive validation, more in-depth analyses of response processes and task-solving strategies in solving CORA tasks were conducted. In particular, we collected students' behavioral data by tracking students' web search histories and eye movements and performed innovative process mining analyses (Schmidt et al., 2020). Additional qualitative analyses of the students' written CORA responses and response processes were conducted using narrative and objective hermeneutic methods (Banerjee et al. 2020). Based on this work, we refined and developed new CORA tasks for the assessment of GEN-COR facets in BRIDGE (for details, see Section 2.4.3).

2.4.2. Domain-specific COR

For the objective, valid assessment of DOM-COR, we used a newly developed computer-based assessment based on a framework initially developed in the Performance Assessment of Learning (PAL) study (Shavelson et al., 2019).[3] The PAL tasks have been developed, tested, and validated for German higher education in accordance with the Standards for Educational and Psychological Testing (AERA et al., 2014). As an example, one of these performance assessment tasks, "Wind Turbine," was developed to assess higher education students' critical reasoning. It consists of a realistic short-frame scenario and incorporates a document library for additional background information on the case including 22 snippets and sources of information that vary in their relevance and trustworthiness. Test takers are assigned a role in a communal real-world decision-making scenario in Germany and are asked to respond by using the given information and to write a policy recommendation for a course of action, i.e., whether or not wind turbines should be built in a small town. The total test time for the PAL task is 60 minutes. The time limit places participants under pressure. It also requires them to decide which sources and arguments to select and review more thoroughly and base this decision on relevance, as there is not enough time to focus on everything provided. Information given in the PAL tasks is classified as being relevant to the problem, trustworthy, or manipulative/lending to errors in judgment/playing to bias.

University students receive higher scores on the PAL tasks, when they, e.g., (i) recognize and use credible information and avoid less credible or peripheral information (facet 'use of sources and arguments'); (ii) avoid judgmental and decision-making 'traps' and biases (facet 'avoidance of, or critical reflection on, heuristics and specific distractor information'); (iii) consider alternative courses of action to the one proposed and indicate why they are giving a certain recommendation; and (iv) display consideration of 'pro'/'con' arguments, including an explanation of different points of view in the task scenario.

[3] Since 2016, in PAL, an international consortium has been focusing on the development and testing of performance assessments as the next generation of measurements of student learning. PAL addresses the question of how performance assessments can enhance targeted student learning beyond rote memorization of facts and actively foster students' acquisition of 21st century skills such as COR (Shavelson et al., 2019).

PAL tasks were validated with university students of different study domains (incl. teaching economics) in accordance with psychometric standards (AERA et al., 2014) (Zlatkin-Troitschanskaia et al., 2019; Nagel et al., 2020b; Zlatkin-Troitschanskaia et al., 2020). For example, in terms of the cognitive validation of the PAL tasks, cognitive interviews were conducted to gather evidence on students' thought processes: Semi-structured interviews with 30 students assessed how the PAL participants integrated and evaluated given information in their decision-making and which individual factors influenced their response processes. The examined questions included, e.g., to what extent the students' answers could be traced back to domain-specific expertise or cross-domain skills and to what extent individual knowledge, pre-conceptions, and beliefs affected the students' information selection and decision-making (for details see Nagel et al., 2020b; Zlatkin-Troitschanskaia et al., 2020). Overall, the results from these validation studies affirm the construct validity and offer evidence of the technical quality of the newly developed PAL task (Shavelson et al., 2019). Therefore, we used this assessment framework to develop DOM-COR tasks in BRIDGE (as described in Section 2.4.3.)

2.4.3. BRIDGE Assessment

Using criterion-sampled performance tasks to assess students' COR in authentic, simulated or even real Internet environments is an innovative approach that has been further developed in BRIDGE to enable the valid measurement of GEN- and DOM-COR skills among young professionals in the three domains. In this context, performance tasks represent complex real-world decision problems and judgment situations (Davey et al., 2015). The realistic scenarios of all tasks used in BRIDGE feature much higher ecological validity compared to prior assessment formats. Most of the studies that previously addressed this issue used traditional research approaches, such as self-reports or simulations, i.e., websites provided in self-contained offline search environments to replicate the appearance of real-world Internet searches (Hargittai et al., 2010). In BRIDGE, by contrast, several performance tasks are being developed in which users work on solving a task by means of real Internet research in a web-based assessment environment, while their handling of online information is captured (log data, event data, and process data).

To measure GEN-COR, in BRIDGE, the CORA tasks have been further developed according to the conceptual framework and its three facets, and then comprehensively validated according to the AERA standards. The developed GEN-COR tasks cover various generic topics and require the participants to check the reliability of certain statements or sources (see Section 2.4.1). No prior content knowledge is required to complete the tasks; participants only need basic abilities to navigate and search the Internet and input and submit their written responses. During the assessment, all web pages the participants access are recorded.

For content and curricular validation, the new CORA tasks developed in BRIDGE were evaluated in individual interviews with twelve experts in the practical education and training phase in the three domains. For cognitive validation, cognitive interviews were conducted with 20 students immediately after solving the CORA tasks, in which they were asked to reflect on their task-solving and reasoning processes. In addition, the collected log data, including the entire browser history with time stamps, were analyzed by means of process mining. The process data analyses indicate that the CORA tasks developed in BRIDGE validly measure the OIA and CIE facets of GEN-COR, whereby the REAS facet is tapped to a lesser extent in these tasks. The Internet sources accessed and information used were additionally rated in terms of their quality, relevance, and topicality, and taken into account when rating participants' GEN-COR skills. After the comprehensive validation, the GEN-COR tasks were administered to young professionals in the three abovementioned domains.

While the same GEN-COR tasks (including the same rating scheme) were used for all three professions, comprehensive domain-specific adaptations were needed for the valid measurement of DOM-COR skills in the different domains. Here, in our first step, we conducted several individual as well as group interviews with our cooperation partners from the practical education and training phases in the three domains. From these interviews, we identified specific scenarios in which the Internet is used particularly frequently as a medium of information during the preparation of practice-related documents. It became apparent, for the teaching profession, that the Internet is used in the preparation of lesson plans, e.g., to prepare up-to-date teaching-learning materials for one's own teaching in a class. In medicine, it is required for preparing a high-quality diagnostic and medical plan using the most current research evidence. In the field of law, when solving practical legal cases, lawyers need to refer to the most current case law, which in some areas is updated daily. Accordingly, for assessing DOM-COR, study participants' entire process of creating the respective documents was recorded and all digital and non-digital information that the study participants consulted and/or used in creating these documents was included in the analyses. The scoring scheme from CORA was adapted to rate this information, e.g., with respect to its grounding in research, research quality, reliability, currentness, and relevance. The scoring scheme from iPAL was adapted for each of the three DOM-COR tasks in the three domains and validated with the help of domain experts for the scoring of the performance data. The texts participants drafted for the DOM-COR tasks were then scored, initially by the project team with respect to DOM-COR skills. In addition, the documents created were evaluated by the instructors from the practical phase with regard to their professional quality. By combining the two assessment and evaluation results for each participant, the relation between DOM-COR and GEN-COR scores and the professional quality of created documents can be determined and facets that can be fostered in practical education are identified.

To summarize, the GEN- and DOM-COR skills assessed via tasks based on the selection of relevant and trustworthy information sources in real online environments have, to our knowledge, not been examined before, neither cross-sectionally nor longitudinally (for assessments that focus on individual COR-relevant facets, see an overview in Molerov et al., 2020). Moreover, using web logging (and eye tracking) to derive response process indicators of how young professionals search for, evaluate, and use information in vivo when solving the GEN- and DOM-COR tasks is a unique approach implemented in this form in BRIDGE, which provides a sound basis for the development of individualized digital trainings to foster these skills in the three domains (see Section 3)

3. DIGITAL TRAINING FRAMEWORK

3.1. INSTRUCTIONAL DESIGN MODEL

Based on the abovementioned prior research, in BRIDGE, we design and implement a web-based learning environment that focuses on the critical handling of online information, which also enables the direct, process-related assessment of these skills (based on log, event, and search histories; see Section 2.4.3). In this context, we develop a new digital COR training concept for the post-academic professional training phase, which consists of various modules with interactive learning tasks that require using the Internet.

In developing a suitable digital training, we focus on one central question: how can the conception of a training for the acquisition and promotion of COR be designed, taking established didactic principles into account? The process of training design follows the principles of the so-called ADDIE instructional design model, which describes the five systematic phases Analyze, Design, Develop, Implement, and Evaluate that are progressively performed in the development of digital

teaching arrangements (Niegemann et al., 2013; Branch, 2009). The first phase, "Analysis," includes identifying and analyzing the needs of the teaching intention and learning offer, defining the overarching guideline and rough learning objectives, as well as determining the target group (Branch, 2009), which together form the basis for the learning content and task formats (Hodell, 2007). In the second phase, "Design," the external structure of the learning offer is determined, e.g., by grouping and sequencing the content into modules. In the third phase, "Develop," the creation and design of the learning materials is performed (Niegemann & Weinberger, 2020), i.e., tasks, including accompanying materials, such as audio tracks and videos, are developed here for each sequence of the training (Hodell, 2007). In the "Implementation" phase, the designed learning offer is tested in practice (Niegemann & Weinberger, 2020). In the final evaluation phase, a summative and/or formative evaluation of the developed learning offer is conducted, which frequently leads to the revision of the concept.

The entire development process of the new training is carried out in iterative loops. Within these individual phases and, in particular, to create effective multimedia materials for learning, further didactic considerations and principles are taken into account, which includes the 12 principles for multimedia learning, such as the "pre-training principle" and "the personalization principle" according to Mayer (2009) (for details, see Kohmer, 2020).

In BRIDGE, within "Analysis," the first phase, the learning environment, the target group, and the contents of the training were analyzed and determined. The needs analysis of the target group revealed that the promotion of COR is hardly or not at all anchored in the curricula of the courses of study in medicine, law, and teaching, although these skills are essential for everyday professional practice (see Sections 2.1; Kohmer, 2020).

Based on the COR construct (see Section 2.2), a competence profile was then developed, which describes the skills to be promoted and based on which the learning content of the training was developed. In the modules, therefore, different abilities are specifically addressed according to the multi-faceted COR concept (for a detailed description of the training conception, see Kohmer, 2020).

In contrast to previous approaches to promoting COR, the distinctive feature of the training is that a complete information research process (according to the IPS-I Modell, Brand-Gruwel & Wopereis, 2006) and associated strategies for obtaining, analyzing, and evaluating reliable online sources are instructed, rather than exclusively assessing the reliability of online sources as in the checklist approach (Christennson, 2006; Blakeslee, 2004; Shanahan, 2008) or the exemplary online media approach (Mathson & Lorenzen, 2008). In this way, the learning units are logically linked and participants already internalize the required steps of the entire COR process during the delivery of the training.

In the second phase, "Design," the content was structured, detailed learning objectives defined, and the time frame of the training was determined. A major advantage of this training is its didactic form of microlearning in short learning units of 15–30 minutes. In this form, the training can be easily and flexibly integrated into the participants' daily lives.

The design will now be followed by the fourth phase, "Implement," in which the training will be technically implemented and practically tested. In the fifth phase, "Evaluate," a revision of the training will be performed (Niegemann & Weinberger, 2020). This not only includes possible improvements but also further developments of the training, e.g., for use on tablets or smartphones or for other target groups, such as in courses during studies or further training in everyday working life.

3.2. FEEDBACK SYSTEM

Closely linked to the issue of training and promoting COR skills is the question of how to provide effective, learning-sensitive feedback, as effective training cannot take place without feedback (Hatziapostolou et al., 2010). The significant relevance of feedback corresponding to a training is undisputed in research: Feedback is considered one of the essential influencing factors of successful learning (Askew, 2004). In this context, feedback is understood to mean any information that learners are offered from an external source, during or after task completion, to confirm or correct errors (Van der Kleij et al., 2015; Hattie & Timperley, 2007; Vasilyeva et al., 2007). It is therefore highly important, both for individual learners and for participating institutions, to receive addressee-appropriate feedback on their performance (Benjes-Small et al., 2013). In this context, in BRIDGE, we focus on the central question: How can an interactive, computer-based, and individualized feedback system be designed, developed, and implemented as a component of a newly developed training program to effectively promote young professionals' GEN- and DOM-COR?

The feedback system being developed in BRIDGE focuses on the newly implemented GEN-COR and DOM-COR assessments, i.e., a holistic performance assessment that maps complex decision problems and assessment situations from the real world (see Sections 2.4.3). Preliminary studies indicate that the performance assessments can also be used for training purposes. Formative assessments are generally considered a core component of effective learning (Bransford et al., 2000). Solving tasks gives learners the opportunity to use their acquired skills to support their learning process (Gikandi et al., 2011). For example, the assessment focuses on specific learning modules, such as 'search strategies' or 'reasoning,' and the participants could be given feedback on the selected COR aspects tapped in these modules.

Providing feedback in educational as well as in more professional contexts involves high demands on time, effort, and appropriate infrastructure (Lachner et al., 2017); particularly written feedback in the form of continuous text or comments on a text entails a great deal of effort for instructors and/or training designers (Mason & Bruning, 2001). This is a particular problem in cases such as GEN- and DOM-COR tasks, where new assignment formats (increasingly) require extensively composed answers, where multiple ways to reach a solution are possible, and where the answer can no longer be selected from several options in a multiple-choice format. In this context, computer-based feedback can be an effective alternative to time-intensive individual feedback by a human evaluator (Ibabe & Jauregizar, 2010). However, this would require a very broad database, which is not yet available in this form.

In the context of such specific requirements and challenges, in BRIDGE, to develop an appropriate feedback system, we follow a model for the design of information systems according to Jacob et al. (2011) and Munzner (2009). According to these models, the following work steps are outlined: 1. task analysis, 2. analysis of information requirements, 3. representation, and 4. implementation. The first step (GEN- and DOM-COR task analysis) serves to define the underlying area of application in detail, i.e. to analyze the specifics of GEN- and DOM-COR in the three domains (Jacob et al., 2011). Only once the underlying areas of application have been clearly defined and elaborated, can aspects relevant to the feedback be considered and incorporated in the development process. The second step (information requirements analysis) entails analyzing and determining the data that are needed for presentation in a feedback system (Jacob et al., 2011). Here, e.g., suitable key figures for use in the context of feedback must be identified, developed, and derived by taking the existing database into account. The third step (representation) deals with decisions regarding the (technical) system used and visual representation. It is important to choose a system that is suitable for the purpose and to observe rules for visualization of feedback. The fourth and last step (implementation) addresses the concrete implementation of the developed

feedback system (for more details, see Lauterbach, 2020). These iterative steps are being implemented in BRIDGE in the context of digital training development.

4. IMPLEMENTATION OF THE WEB-BASED ASSESSMENT AND LEARNING ENVIRONMENT IN BRIDGE

Since the assessment and training (including feedback) are digital, one of the central questions is how it should be designed in technical terms. The actual technical implementation depends on the specific (didactic) design of the individual assessments, training modules, and assessment-based feedback in BRIDGE as described above. According to this design model, by developing the new training (with feedback system) in BRIDGE, we use modern information technologies, which offer numerous possibilities for creating multimedia learning materials and content based on various sources and integrating them into a digital training environment (Dabbagh & Kitsantas, 2012; van Merriënboer, 2002), i.e., a learning management system (LMS) that provides BRIDGE participants with access to computer-based testing, training modules, and feedback. By creating this new BRIDGE LMS platform using the GEN- and DOM-COR tasks and digital learning modules, participants are enabled to independently and self-directedly assess and train their COR skills.

A new web-based environment, specifically used as an assessment, training, and feedback platform in BRIDGE, is launched from the cloud-based software Microsoft Azure. Participants can access the assessment from home using a computer/laptop via remote desktop connection they have been provided with for this purpose. They are presented with the GEN- and DOM-COR tasks via this assessment platform. The participants solve the tasks by conducting a real Internet search, while implemented tracking programs record their browsing history in the background and then store the data in a database. The data generated in this way forms the basis for the feedback environment. Custom programs, like the BRIDGEConnector, an independently developed program that generates log data of the browsing history and uploads it to a database, were developed in the design stages of the assessment environment to enable the process of data generation and storage. For a more detailed description of the individual elements involved in the assessment environment and process and data generation, see Lauterbach (2020).

For the development and implementation of the web-based assessment and learning environment, a very central aspect was to ensure data protection compliance and research ethics according to the applicable German law, including the German General Data Protection Regulation (GDPR) in BRIDGE. To this end, the project-specific data management plan was developed and approved.

Since the training is designed as Web-Based Training (WBT), it can be flexibly used regardless of location and does not require physical presence. Another advantage of WBT is that the content can also be flexibly adapted anytime (Stoecker, 2013). This is particularly important as this is the initial design of the BRIDGE training, and it can be expected that improvements will need to be made after practical testing and revision. For this reason, in phase three, "Develop," care has been taken in the design of the user interface and the individual learning units to ensure that users can always provide feedback on the extent to which content should be added or modified. Central to this are the accompanying pre- and post-assessments in BRIDGE (see Section 5). To this end, after the first GEN-COR und DOM-COR assessment (described above), in an approximately six-week online training course, participants are instructed and supported in the general and domain-specific use of online media and effective strategies for the competent handling of information in both general and professional contexts. After the digital training intervention, participants will be retested while performing GEN- and DOM-COR tasks. Testing participants before and after the training (see Figure 2 in the Supplement for project timetable) provides a differentiated insight into the effectiveness of the newly developed online training in terms of changes in the participants' general and domain-specific use of online media (i.e., GEN- and DOM-COR). In addition, the

accompanying reflection tasks indicate to both the participants and the training developers the extent to which the content was successfully conveyed.

5. RESEARCH DESIGN, SAMPLE, AND ANALYSES

To achieve the research objectives and to test the hypotheses, the three-year project measures young professionals' media use in their 1-to-2-year practical training phases. The sample consists of medical, legal, and teaching (subject: economics) trainees in their practical phase from the German federal states Rhineland-Palatinate and Hesse, who are undergoing a professional training cycle following their university studies, beginning in November 2020 or later, so that the project includes several cohorts who enter their practical phase in 2020 or 2021. Thus, the group of participants consists of young professionals working in the three different professional areas and who participate in this assessment and digital training on a part-time and voluntary basis.
These three groups have similar training structures (Fabry, 2016) and are therefore suitable for a comparative analysis with regard to the structural characteristics of the professional documents they prepare (legal and medical case reports and teaching plans). For the three professions, the importance of the use of online information in the creation of such documents is particularly evident (see Section 2.1). At the same time, these professions differentiate domain-specifically, e.g., with regard to manner and amount of learning offerings and opportunities in the practical phase (Steffens et al., 2017) and, thus, allow the investigation of the domain-specificity of the development of professional media use in these different professional educational contexts. Accordingly, the development of skills in the job-related use of online information is captured via the constructs GEN- and DOM-COR and compared between the three domains.
To enable the participation in all online assessments and trainings, study participants are equipped with prepared laptops with tracking programs (as recommended in prior research) to assess media use in professional practice in ecologically valid settings (Steffens et al., 2017; Brooks, 2016).
The surveys consist of the following steps and phases:
1. Initial survey at the beginning of practical training: Social demographics (e.g., gender, age) and the participants' general use of online information (GEN-COR) using CORA tasks are initially collected in an online survey of job entrants from the three domains. Based on the distribution of these personal characteristics and the participants' GEN-COR levels, approx. N = 30 job entrants per cohort and per domain are selected to create a purposeful sample (Palinkas et al., 2015).
2. Pre-testing: In a first DOM-COR performance assessment, the job entrants are asked to create real practical professional documents (lesson plans, medical/legal case reports) using the prepared laptops and digital media when preparing these documents. In this way, their use of online information is assessed in a direct and valid way and all online sources accessed (web pages with videos, images, texts, etc.) are directly recorded. The use of nondigital resources is assessed via self-reporting during the solving of the DOM-COR task.
3. Online training: In an approx. 4-to-6-week online training, the study participants' GEN-COR and DOM-COR skills are fostered by means of digital teaching-learning packages with feedback systems, which address different COR facets.
4. Post-testing: After the online training and using the same format as in the pre-test, the development of the participants' GEN-COR and the development of their domain-specific use of online information (DOM-COR) will be assessed again.
5. Effectiveness analysis: For every cohort and domain, longitudinal evaluations of the captured process-related data (e.g., log files) and performance data (e.g. COR scores), taking into account several external criteria, such as the trainee's final grades at the end of their practical training, provide information on the effectiveness of the newly implemented digital training for the development of GEN- and DOM-COR in practical professional education contexts and on how

the interaction between the informal and formal teaching-learning opportunities during the practical phase of professional education can be optimized.

The process- and performance-related analysis of the development of GEN- and DOM-COR and how it can be effectively supported in professional vocational educational processes is based on a systematic integration of several internationally established approaches, which stem from a range of superordinate theories, such as the theories of domain learning (Alexander, 2003) and multimedia learning (Mayer, 2009), theories on the handling of (multimodally represented) information from multiple texts (Braasch & Bråten, 2017; List & Alexander. 2017; Wiley et al., 2009; Yu et al., 2018), Media Literacy Research (Steffens et al., 2017) as well as process-related research on educational/learning behavior based on verbal (Brückner & Pellegrino, 2016; Leighton, 2017; Ercikan & Pellegrino, 2017; Zumbo & Hubley, 2017) and computer-generated data (such as log data, eye movement, page dwell times) (Goldhammer & Zehner, 2017; Li et al., 2017; Oranje et al., 2017; Russell & Huber, 2017).

The performance data, i.e., the written professional documents created by the study participants (i.e., lesson plans, legal and medical case reports), are evaluated using a validated scoring scheme (Zlatkin-Troitschanskaia et al., 2019) in BRIDGE (see Section 2.4.3). Moreover, these documents will be evaluated by practice partners (instructors) from the three domains and the results of both evaluations will be compared, reflected, and used to build the foundation for a detailed feedback for the participants (see Section 3.2).

To analyze the process data from the collected log data, e.g., browsing patterns and the online sources used are examined. Analyses of learning process data provide nuanced insights into how specific information is used to accomplish job-related tasks. In BRIDGE, therefore, the use of online information is measured during the creation of real practical professional documents via both the analysis of these created documents and their quality in relation to the captured process data (e.g., log files with timestamps). In addition, the texts from web pages used by participants will be analyzed with computational-linguistic methods and the correlations between the textual analyses and the findings from learning data will be statistically integrated in a mixed method design. Automated analysis methods allow for the integration of different types of data, such as the documents created by the study participants and the process data collected during the drafting process (e.g., eye movements while searching for information, time spent on web pages, number of web pages visited). After every survey and assessment, the key results will be reported back to the participants in the form of computer-based feedback.

To analyze the DOM-COR and GEN-COR skills that participants demonstrate in these performance assessments, innovative approaches such as data mining and text mining are also used to, for instance, (semi-automatically) evaluate the documents written by the study participants; approaches from the field of learning analytics are used for generating automated feedback.

This comprehensive and in-depth analysis of process and performance data before and after the training also provides evidence in terms of the effectiveness of this training and feedback system and how both can be further developed and optimized for a broad and successful transfer into practice to promote COR skills among young professionals.

Another particular focus of analysis lies on the investigation of domain-specificity, i.e., the question to what extent the development of COR skills can be considered domain-specific or generic. The answer to this question has significant implications for the creation of future training approaches within and/or across domains. To this end, the level and the development of DOM- and GEN-COR skills are comparatively considered in three domains using the method of known groups (Hubley & Zumbo, 2011).

6. CONCLUSION

The importance of COR and of the corresponding assessment, training and feedback as measures to effectively enhance this essential skill is central to BRIDGE. To this end, BRIDGE develops a digital training approach to promote GEN- and DOM-COR. In this context, BRIDGE examines the critical use of online media and how it can be effectively promoted in educational processes in professional training phases among medical interns and prospective teachers/law clerks in two German federal states. BRIDGE empirically implements new theoretical-conceptual modeling approaches and online assessments to validly measure GEN- and DOM-COR skills. Their development over the practical training phase is analyzed in an innovative MTMM research design, using a longitudinal approach. Based on these findings and the corresponding feedback, digital training tools like those newly developed and implemented in BRIDGE can be used more effectively in practice. Thanks to the continuous collaboration with partners from educational practice (e.g., trainers) during the traineeship (law/teaching) and practical year (medicine), it can be ensured that experts are involved in the subject-specific qualitative analyses and evaluations and project results are transferred into practice.

In this way, BRIDGE makes a significant contribution to the objectives of the new research program (see Section 1.1) by implementing new assessment environments, developing new digital training and feedback tools, and providing internationally unique findings on (i) the development of young professionals' skills regarding media use and the effective promotion of these skills using the new digital training opportunities as implemented in BRIDGE; (ii) the relations between the development of these skills and the completion of formal learning opportunities in the practical professional training phase in the three domains; and (iii) the implementation of new digital training solutions for effectively promoting professional media use and the transfer thereof into professional practice, e.g., using automated evaluation and feedback systems.

REFERENCES

Abreu, B. S. de, Mihailidis, P., Lee, A. Y. L., Melki, J., & McDougall, J. (Eds.). (2017). *International handbook of media literacy education*. New York, London: Routledge Taylor & Francis Group.

AERA, APA & NCME (2014). *Standards for educational and psychological testing.* AERA.

Ainley J., Fraillon, J. Schulz, W. & Gebhardt, E. (2016). Conceptualizing and Measuring Computer and Information Literacy in Cross-National Contexts. *Applied Measurement in Education, 29*(4), 291 – 309. https://doi.org/10.1080/08957347.2016.1209205

Alexander, P. A., Jablansky, S., Singer, L. M., & Dumas, D. (2016). Relational reasoning: What we know and why it matters. *Policy insights from the behavioral and brain sciences*, *3*(1), 36-44. https://doi.org/10.1177/2372732215622029

Alexander, P. A. (2003). The Development of Expertise: The Journey From Acclimation to Proficiency. *Educational Researcher*, *32*(8), 10-14. https://doi.org/10.3102/0013189X032008010

Allen, D. & Harkins, K.J. (2005). Too much guidance? *The Lancet, 365*(9473), 21-27.

Amin, J. (2016). Redefining the Role of Teachers in the Digital Era. *International Journal of Indian Psychology*, *3*(3). https://doi.org/10.25215/0303.101

Askew, S. (2004). *Feedback for Learning.* Routledge.

Banerjee, M., Zlatkin-Troitschanskaia, O., & Roeper, J. (2020). Narratives and Their Impact on Students' Information Seeking and Critical Online Reasoning in Higher Education Economics and Medicine. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.570625

Basak, D., & Schimmel, R. (2008). Internet im Jurastudium – Plädoyer für einen wohlüberlegten Einsatz des WWW. *Zeitschrift für das juristische Studium,* 4(94), 435-440.

Benjes-Small, C., Archer, A., Tucker, K., Vassady, L. & Resor, J. (2013). Teaching Web Evaluation. *Communications in Information Literacy, 7*(1), 39-49.

Blakeslee, S. (2004). The CRAAP Test. LOEX Quarterly, 31(3), 6-7.

Bundesministerium für Bildung und Forschung (BMBF) (2015). *Technologiebasierte Kompetenzmessung in der beruflichen Bildung (ASCOT) Ergebnisse und Bedeutung für Politik und Praxis.* BMBF.

Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) Model: Basic Assumptions and Preliminary Evidence. *Educational Psychologist*, *52*(3), 167-181. https://doi.org/10.1080/00461520.2017.1323219

Branch, R.M. (2009). Instructional Design: The ADDIE-Approach. New York: Springer

Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., & van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning*, *33*(3), 234-251. https://doi.org/10.1111/jcal.12162

Brand-Gruwel, S. & Wopereis, I. (2006). Integration of the information problem-solving skill in educational programme: The effects of learning with authentic tasks. *Technology, Instruction, Cognition, and Learning*, 4, 243-263.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School.* National Academy Press.

Brooks, C. (2016). *ECAR study of students and Information Technology.* ECAR.

Brückner, S., & Pellegrino, J. W. (2016). Integrating the Analysis of Mental Operations Into Multilevel Models to Validate an Assessment of Higher Education Students' Competency in Business and Economics. *Journal of Educational Measurement*, *53*(3), 293-312. https://doi.org/10.1111/jedm.12113

Booth, C. (2011). Reflective teaching, effective learning: Instructional literacy for library educators. Chicago: American Library Association.

Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105.

Chiu, Y.-L., Liang, Y.-C., & Tsai, C.-C. (2013). Internet-specific epistemic beliefs and self-regulated learning in online academic information searching. *Metacogn. Learn.* 8, 235–260. doi: 10.1007/s11409-013-9103-x

Christennson, K. (2006). Radcab: Your vehicle for Information Evaluation. Fort Atkinson: Highsmith Inc.

Dabbagh, N. & Kitsantas, A. (2012). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education*, *15*(1), 3-8.

Davey, T., Ferrara, S., Shavelson, R., Holland, P., Webb, N., & Wise, L. (2015). Psychometric considerations for the next generation of performance assessment. Washington, DC: Center for K-12 Assessment & Performance Management, Educational Testing Service. Retrieved from https://www.ets.org/Media/Research/pdf/psychometric_considerations_white_paper.pdf

Ercikan, K., & Pellegrino, J. W. (Eds.). (2017). *NCME applications of educational measurement and assessment book series. Validation of score meaning for the next generation of assessments: The use of response processes.* Routledge.

Fabry, G. (2016). Warum Hochschuldidaktik? Die Perspektive der Humanmedizin. *Zeitschrift Für Didaktik Der Rechtswissenschaft*, *3*(2), 136-151. https://doi.org/10.5771/2196-7261-2016-2-136

Facione, P. (1990). *The Delphi Report: Executive Summary; Critical thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*. The California Academic Press.

Fischer, F. (Ed.). (2018). *Scientific reasoning and argumentation: The roles of domain-specific and domain-general knowledge*. New York NY: Routledge.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4). 25–42. https://doi.org/10.1257/089533005775196732

Gadiraju, U., Yu, R., Dietze, S., & Holtz, P. (2018). Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web. In C. Shah, N. J. Belkin, K. Byström, J. Huang, & F. Scholer (Eds.), *CHIIR'18: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*. (pp. 2-11). The Association for Computing Machinery. https://doi.org/10.1145/3176349.3176381

Gikandi, J.W., Morrow, D. & Davis, N.E. (2011). Online Formative Assessment in Higher Education: A Review of the Literature. *Computers and Education*, 57, 2333-2351. http://dx.doi.org/10.1016/j.compedu.2011.06.004

Goldhammer, F., & Zehner, F. (2017). What to Make Of and How to Interpret Process Data. *Measurement: Interdisciplinary Research & Perspective*, *15*(3-4), 128-132. https://doi.org/10.1080/15366367.2017.1411651

Hague, C., and Payton, S. (2010). *Digital Literacy Across the Curriculum*. Futurelab

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication*(4), 468-494.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research. 77*(1), 81–112.

Hatziapostolou, T. & Paraskakis, I. (2010). Enhancing the Impact of Formative Feedback on Student Learning through an Online Feedback System. *Electronic Journal of e-Learning, 8*(2), 111-122.

Hocevar, K. P., Flanagin, A. J., & Metzger, M. J. (2014). Social media self-efficacy and information evaluation online. *Computers in Human Behavior*, *39*, 254-262. https://doi.org/10.1016/j.chb.2014.07.020

Hölscher, C., & Strube, G. (2000). Web search behaviour of Internet experts and newbies. *Computing Networks*, *33*(1), 1–6. https://doi.org/10.1016/S1389-1286(00)00031-1

Hodell, C. (2007). Basics of Instructional Systems Development. Alexandria: ASTD.

Hubley, A. M., & Zumbo, B. D. (2011). Validity and the Consequences of Test Interpretation and Use. *Social Indicators Research*, *103*(2), 219-230. https://doi.org/10.1007/s11205-011-9843-4

Jacob, O., Weiß, N. & Schweig, J. (2011). *Konzeption und Gestaltung von Management Dashboards*. Working Paper, Nr. 18, Hochschule für Angewandte Wissenschaften Neu-Ulm.

Jahn, D., & Kenner, A. (2018). Critical Thinking in Higher Education: How to foster it using Digital Media. In D. Kergel, B. Heidkamp, P. K. Telléus, T. Rachwal, & S. Nowakowski (Eds.), *The Digital Turn in Higher Education* (pp. 81–109). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19925-8_7

Ibabe, I. & Jauregizar, J. (2010). Online self-assessment with feedback and metacognitive knowledge. *Higher Education*, *59*, 243–258

Kimmerle, J., Moskaliuk, J., Oeberst, A. & Cress, U. (2015). Learning and Collective Knowledge Construction With Social Media: A Process-Oriented Perspective. *Educational Psychologist*, *50*(2), 120-137. https://doi.org/10.1080/00461520.2015.1036273

Kohmer, A. (2020). Entwicklung und Validierung eines Trainings zur Erfassung und Förderung des kritischen Umgangs mit Online-Medien. Masterthesis.

Kuhn, S., Müller, N., Kirchgässer, E., Ulzheimer, L. & Lucia Deutsch, K. (2020). Digital skills for medical students – qualitative evaluation of the curriculum 4.0 "Medicine in the digital age". *Journal for Medical Education*, *37*(6), Doc60. https://doi.org/10.3205/zma001353

Lachner, A., Burkhart, C. & Nückles, M. (2017). Formative computer-based feedback in the university classroom: Specific concept maps scaffold students' writing. *Computers in Human Behavior*, *72*, 459–469.

Lauterbach, B. (2020). *Konzeption und Entwicklung eines Feedbacksystems als Teil eines digitalen Lehr-Lernarrangements zur Förderung von Critical Online Reasoning (COR).* Masterthesis.

Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research*. *Understanding qualitative research*. Oxford University Press

Li, Z., Banerjee, J., & Zumbo, B. D. (2017). Response Time Data as Validity Evidence: Has it lived up to its promise and, if not, what would it take to do so. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and Investigating Response Processes in Validation Research* (pp. 159-178). Springer International Publishing.

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R*. Hogrefe.

List, A., & Alexander, P. A. (2017). Analyzing and Integrating Models of Multiple Text Comprehension. *Educational Psychologist*, *52*(3), 143-147. https://doi.org/10.1080/00461520.2017.1328309

Maireder, A. & Nagl, M. (2010). Internet in der Schule, Schule im Internet. Schulische Kommunikationskultur in der Informationsgesellschaft. *In mediamanual.* Texte 2010, Nr.1. http://www2.mediamanual.at.

Mason, B. J. & Bruning, R. H. (2001). *Providing feedback in computer-based instruction: What the research tells us.* Center for Instructional Innovation.

Mathson, S. M. & Lorenzen, M. G. (2008). We won't be fooled again: teaching critical thinking via evaluation of hoax and historical revisionist Websites in a library credit course. *College & Undergraduate Libraries*, 15(1/2), 211-230.

Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O. & Jitomirski, J. (2020). Positive and Negative Media Effects on University Students' Learning: Preliminary Findings and a Research Program. In O. Zlatkin-Troitschanskaia (Eds.), *Frontiers and Advances in Positive Learning in the Age of Information* (pp. 109–119). Springer. https://doi.org/10.1007/978-3-030-26578-6_8

Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge, England: Cambridge University Press.

Mathson, S. M. & Lorenzen, M. G. (2008). We won't be fooled again: teaching critical thinking via evaluation of hoax and historical revisionist Websites in a library credit course. *College & Undergraduate Libraries*, *15*(1/2), 211-230.

Meredith, S. (2010). First year law students, legal research skills & electronic resources. *The Law Teacher*, *41*(2), 191-205. https://doi.org/10.1080/03069400.2007.9959738

McGrew, S., Smith, M., Breakstone, J., Ortega, T. & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology, 89*(3), 485-500. https://doi.org/10.1111/bjep.12279

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory and Research in Social Education, 46*(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

McGrew, S., Ortega, T., Breakstone, J., & Wineburg, S. (2017). The challenge that's bigger than fake news. Civic reasoning in a social media environment. *American Educator, 41*(9), 4–9.

Mcquail, D. (1991). Media Performance Assessment in the Public Interest: Principles and Methods. *Annals of the International Communication Association*, *14*(1), 111-145. https://doi.org/10.1080/23808985.1991.11678782

Mesko, B. & Győrffy, Z. (2019). The Rise of the Empowered Physician in the Digital Health Era: Viewpoint. Journal of Medical Internet Research, 21(3), e12490.

Merriënboer, J. J. G. van, Clark, R. E. & Croock, M. B. M. de (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology Research and Development*, *50*(2), 39-61.

Mielke, B. & Wolff, C. (2012). Ausbildungskonzepte zur Verbesserung juristischer Informationskompetenz. Vortrag präsentiert auf der IRIS Konferenz, Salzburg. https://www.researchgate.net/publication/236340879_Ausbildungskonzepze_zur_Verbesserung_juristischer_Informationskompetenz.

Molerov, D., Zlatkin-Troitschanskaia, O., Nagel, M.T., Brückner, S., Schmidt, S. & Shavelson, R. (2020). Assessing University Students' Critical Online Reasoning Ability: A Conceptual and Assessment Framework with Preliminary Evidence. *Frontiers in Education, 5*(1), 1-29. https://doi.org/10.3389/feduc.2020.577843

Molerov, D., Zlatkin-Troitschanskaia, O., and Schmidt, S. (2019). Adapting the civic online reasoning assessment cross-nationally using an explicit functional equivalence approach. In *Annual Meeting of the American Educational Research Association (Toronto).*

Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 921-928. https://doi.org/10.1109/TVCG.2009.111.

Nagel, M.-T., Schäfer, S., Zlatkin-Troitschanskaia, O., Schemer, C., Maurer, M. & Molerov, D. (2020a). How do university students' web search behavior, website characteristics, and the interaction of both influence students' critical online reasoning? *Frontiers in Education*, *5*(1). https://doi.org/ 10.3389/feduc.2020.565062

Nagel, M.-T., Zlatkin-Troitschanskaia, O., Schmidt, S., & Beck, K. (2020b). Performance Assessment of Generic and Domain-Specific Skills in Higher Education Economics. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper & C. Lautenbach (Eds.), *Student Learning in German Higher Education: Innovative Measurement Approaches and Research Results* (p. 281–299). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-27886-1_14

National Research Council (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. National Academies Press

Niegemann, H. & Weinberger, A. (2020). Handbuch Bildungstechnologie: Konzeption und Einsatz digitaler Lernumgebungen. Wiesbaden: Springer.

Niegemann, H., Hessel, S., Hoschscheid-Mauel, D., Aslanski, K. & Deimann, M., Kreuzberger, G. (2013). Kompendium E-Learning. Berlin: Springer.

O'Carroll, A.M., Westby, E.P., Dooley, J. & Gordon, K.E. (2015). Information-Seeking Behaviors of Medical Students: A Cross-Sectional Web-Based Survey. *Journal of Medical Internet Research Medical Education, 1*(1), 2

Oranje, A., Gorin, J., Jia, Y., & Kerr, D. (2017). Collecting, Analyzing, and Interpreting Response Time, Eye-Tracking, and Log Data. In K. Ercikan & J. W. Pellegrino (Eds.), *NCME applications of educational measurement and assessment book series. Validation of score meaning for the next generation of assessments: The use of response processes* (pp. 39-51). Routledge

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N. & Hoagwood, K. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and policy in mental health*, *42*(5), 533-544. https://doi.org/10.1007/s10488-013-0528-y

Paul, R. & Elder, L. (2005). *A Guide for Educators to Critical Thinking Competency Standards, Principles, Performance Indicators, and Outcomes with a Critical Thinking Master Rubric*. Foundation for Critical Thinking.

Redecker, C. (2017*). European Framework for the Digital Competence of Education: DigCompEdu*. Luxembourg Publications Office of the European Union. https://doi.org/10.2760/159770

Reichert, F., Zhang, D., Law, N.W.Y., Wong, G.K.W., de la Torre, J. (2020). Exploring the structure of digital literacy competence assessed using authentic software applications. *Educational Technology Research and Development, 68*, 2991–3013.

Rott, K.J. (2014). Medienkompetenz im Studium: Wie gut ist die Vorbereitung für das spätere Berufsfeld? In O. Zawacki-Richter, D. Kergel, N. Kleinefeld, P. Muckel, J. Stöter & K. Brinkmann (Hrsg.), Teaching Trends 2014. Offen für neue Wege: Digitale Medien in der Hochschule Münster (S.153-169). New York: Waxmann.

Russell, L. B., & Huber, U. (2017). Some Thoughts on Gathering Response Process Validity Evidence: in the Context in Online Measurement and Digital Revolution. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and Investigating Response Processes in Validation Research* (pp. 229-250). Springer International Publishing.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91*(3), 497–510. https://doi.org/10.1037/0022-0663.91.3.497

Schiefner-Rohs, M. (2012). Verankerung von medienpädagogischer Kompetenz in der universitären Lehrerbildung. In R. Schulz-Zander, R.B. Eickelmann, H. Moser, H. Niesyto & P.Grell (Hrsg.), *Jahrbuch Medienpädagogik 9*.,(S.359-387). Wiesbaden: Springer.

Schimmel, R. (2011). *Recherche im Jurastudium. Bessere Noten mit besseren Suchmaschinen-Strategien.* https://www.lto.de/recht/studium-referendariat/s/recherche-im-jurastudium-bessere-noten-mit-besseren-suchmaschinen-strategien/

Schmidt, S., Zlatkin-Troitschanskaia, O., Roeper, J., Klose, V., Weber, M., Bültmann, A.-K., & Brückner, S. (2020). Undergraduate Students' Critical Online Reasoning: Process Mining Analysis. *Frontiers in Psychology.* Advance online publication. https://doi.org/10.3389/fpsyg.2020.576273

Senkbeil, M. (2018). Development and validation of the ICT motivation scale for young adolescents. Results of the international school assessment study ICILS 2013 in Germany. *Learning and Individual Differences*, *67*, 167-176. https://doi.org/10.1016/j.lindif.2018.08.007

Shanahan, M.C. (2008). Transforming information search and evaluation practices of undergraduate students. *International Journal of Medical Information*, 77(8), 518-526.

Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., & Marino, J. P. (2019). Assessment of University students' critical thinking: next generation performance assessment. *International Journal of Testing, 19*(4), 337–362. https://doi.org/10.1080/15305058.2018.1543309

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*, 117 – 143. https://doi.org/10.1037/pspp0000096

Steffens, Y., Schmitt, I. L., & Aßmann, S. (2017). *Mediennutzung Studierender: Über den Umgang mit Medien in hochschulischen Kontexten. Systematisches Review nationaler und internationaler Studien zur* Mediennutzung Studierender. https://doi.org/10.13154/rub.106.95

Stoecker, D. (2013). *eLearning-Konzept und Drehbuch: Handbuch für Medienautoren und Projektleiter*. Springer-Verlag.

Taylor, A., & Dalal, H. A. (2014). Information literacy standards and the World Wide Web: results from a student survey on evaluation of Internet information sources. *Information Research*, 19(4).

Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 94*(1), 197–209. https://doi.org/10.1037/0022-0663.94.1.197

Tossell, C. C., Kortum, P., Shepard, C., Rahmati, A., & Zhong, L. (2015). You can lead a horse to water but you cannot make him learn: Smartphone use in higher education. *British Journal of Educational Technology*, *46*(4), 713-724. https://doi.org/10.1111/bjet.12176

Urban, J., & Schweiger, W. (2013). News quality from the recipients' perspective. *Journalism Studies*, *15*, 821–840. doi: 10.1080/1461670X.2013.856670

Van der Kleij, F. M., Feskens, R. C. & Eggen T. J. (2015). Effects of Feedback in a Computer-based Learning Environment on Students Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, *85*(4), 475–511.

Vasilyeva, E., Puuronen, S., Pechenizkiy, M. & Räsänen, P. (2007). Feedback adaption in web-based learning systems. *International journal of continuing Engineering Education and Life Long Learning, 17*, 337–357.

Wagner, J. (2018). *Legal Tech und Legal Robots. essentials.* Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20057-2

Walton, D. (2006a). *Fundamentals of Critical Argumentation. Critical Reasoning and Argumentation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511807039

Walton, D. (2006b). Rules for Reasoning from Knowledge and Lack of Knowledge. *Philosophia*, *34*, 355 – 376.

Watson, H. & Burr, S. (2018). Research skills in medical education. MedEdPublish, 7(3). https://doi.org/10.15694/mep.2018.0000151.1

Weber, H., Becker, D. & Hillmert, S. (2019). Information-seeking behaviour and academic success in higher education: Which search strategies matter for grade differences among university students and how does this relevance differ by field of study?. *Higher Education, 77*(4), 657–678. https://doi.org/10.1007/s10734-018-0296-4

Weber, H., Hillmert, S., & Rott, K. (2018). Can digital information literacy among undergraduates be improved? Evidence from an experimental study. *Teaching in Higher Education, 23*(8), 909-926. https://doi.org/10.1080/13562517.2018.1449740.

White, R., Dumais, S. & Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. WSDM '09: *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, 132-141. https://doi.org/10.1145/1498759.1498819

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source Evaluation, Comprehension, and Learning in Internet Science Inquiry Tasks. *American Educational Research Journal*, *46*(4), 1060-1106. https://doi.org/10.3102/0002831209333183

Wineburg, S. & McGrew, S. (2018). Lateral Reading and the Nature of Expertise: Reading Less and Learning More When Evaluating Digital Information. Stanford History Education Group Working Paper, 2017(A1). http://dx.doi.org/10.2139/ssrn.3048994

Wineburg, S., Breakstone, J., McGrew, S. & Ortega, T. (2018). Why google can't save us. The challenges of our post-gutenberg moment. In O. Zlatkin-Troitschanskaia, G. Wittum, & A. Dengel (Eds.), *Positive Learning in the Age of Information* (pp.221–228). Springer. https://doi.org//10.1007/978-3-658-19567-0_13

Yu, R., Gadiraju, U., Holtz, P., Rokicki, M., Kemkes, P., & Dietze, S. (Eds.) (2018). *Predicting User Knowledge Gain in Informational Search Sessions*. http://arxiv.org/pdf/1805.00823v1

Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., & Krstev, J. (2021, in press). Students' Online Information Use and Learning Progress in Higher Education – A Critical Literature Review. *Studies in Higher Education. Special Issue.*

Zlatkin-Troitschanskaia, O., Beck, K., Fischer, J., Braunheim, D., Schmidt, S. & Shavelson, R. J. (2020). The role of students' beliefs when critically reasoning from multiple contradictory sources of information in performance assessments. *Frontiers in Education, 11*(2192). https://doi.org/10.3389/fpsyg.2020.02192

Zlatkin-Troitschanskaia, O., Shavelson, R. J., Schmidt, S., & Beck, K. (2019). On the complementarity of holistic and analytic approaches to performance assessment scoring. *The British Journal of Educational Psychology,* 89(3), 468–484. https://doi.org/10.1111/bjep.12286

Zlatkin-Troitschanskaia, O., Toepper, M., Molerov, D., Buske, R., Brückner, S., Pant, H. A., Hofmann, S., & Hansen-Schirra, S. (2018). Adapting and Validating the Collegiate Learning Assessment to Measure Generic Academic Skills of Students in Germany: Implications for International Assessment Studies in Higher Education. In O. Zlatkin-Troitschanskaia, M. Toepper, H. A. Pant, C. Lautenbach, C. Kuhn (Eds.), *Assessment of Learning Outcomes in higher education – Cross-National Comparisons and Perspectives* (pp. 245-266). Springer.

Zumbo & A. M. Hubley (Eds.), *Understanding and Investigating Response Processes in Validation Research*. Springer International PublishingSUPPLEMENT

Figure 1. **Exemplary CORA task.**

TASK 5: Government revenue

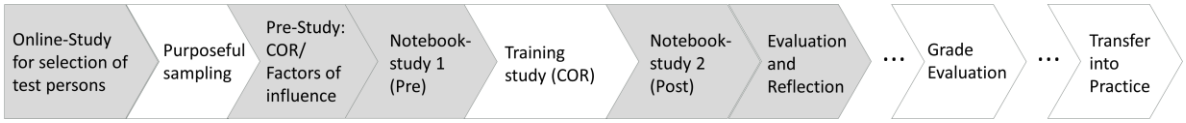
The following tweet appears in your Twitter feed: https://twitter.com/XXXX

To what extent is the tweet suitable as a source of information on government revenues in recent years? Answer both questions.

1. Why could this tweet be a reliable source of information on government revenue in recent years?
2. Why could this tweet be an unreliable source of information on government revenue in recent years?

You can use any information on this website, as well as search the Internet. Justify your answers with evidence from the Internet sources used and include the corresponding URLs.

Figure 2. **Research schedule.**



ABOUT THE AUTHORS

Olga Zlatkin-Troitschanskaia
Olga Zlatkin-Troitschanskaia has been Chair of Business and Economics Education at Johannes Gutenberg University Mainz (JGU), Germany, since 2006. She earned her doctoral degree from Humboldt University of Berlin in 2004 and her postdoctoral qualification in 2006. She has published widely on empirical educational research in vocational and higher education. She has directed numerous externally funded national and international research projects and has been coordinating the national research program 'Modeling and Measuring Competencies in Higher Education (KoKoHs)' since 2011; she also co-implemented the international collaborative research project 'Performance Assessment of Learning' (iPAL). Her research has earned various awards and honors. She is a member of many national and international research academies as well as advisory and editorial boards, and serves as an expert consultant to ministries, foundations, and academic journals.

Sebastian Brückner
Dr. Sebastian Brückner is a postdoctoral researcher at the Chair of Business and Economics Education at the Johannes Gutenberg University (JGU) in Mainz, Germany. In 2015 he received his doctoral degree from the JGU and has earned several research awards. His research focuses on modeling and assessment of teaching and learning in digital and analog environments in the domain of business and economics using technological methods (e.g., eye-tracking).

Marie-Theres Nagel
Marie-Theres Nagel is a research assistant at the Chair of Business and Economics Education at the Johannes Gutenberg University (JGU) in Mainz, Germany. Since 2017 she has been pursuing her doctorate with two projects. The first project is a BMBF-funded project on the valid assessment of economics knowledge and its development in business and economics university students, while the second project focuses on the valid assessment of critical thinking skills (in online contexts) of students in higher education. Additionally, since 2020, she has been working in a BMBF-funded project on young professionals' development of critical online thinking, called "Vocational education processes in law, teaching and medicine using digital media (BRIDGE)".

Ann-Kathrin Bültmann
Ann-Kathrin Bültmann a research assistant at the Chair of Business and Economics at the Johannes Gutenberg University (JGU) in Mainz, Germany. She studied business education at Johannes Gutenberg University (JGU) and has been working as a student assistant at the Chair of Business and Economics Education at JGU since 2016 in various projects such as "Positive Learning in the Age of InformaTiOn (PLATO)". Since 2020, she has been working in a BMBF-funded project on young professionals' development of critical online thinking, called "Vocational education processes in law, teaching and medicine using digital media (BRIDGE)".

Jennifer Fischer
Jennifer Fischer is a research assistant at the Chair of Business and Economics Education at the Johannes Gutenberg University (JGU) in Mainz, Germany and currently pursuing her doctorate. She studied business education at the Johannes Gutenberg University (JGU). Since 2018, she has worked in the research projects "Modeling and Measuring Competencies in Higher Education (KoKoHs)", "Study Success and Opportunities for Refugees in Higher Education - An Effectiveness Analysis (SUCCESS)" and "Positive Learning in the Age of Information (PLATO)". Currently, she is working in a BMBF-funded project on young professionals' development of

critical online thinking, called "Vocational education processes in law, teaching and medicine using digital media (BRIDGE)".

Susanne Schmidt
Dr. Susanne Schmidt is a research associate and postdoctoral researcher at the Chair of Business and Economics Education at Johannes Gutenberg Mainz University (Germany), where she earned her doctoral degree in 2016 and has received several research awards. Her research focus lies on analyzing longitudinal studies as well as learning analytics (e.g., analyzing process data from digital learning environments) to describe the learning processes of higher education students and graduates.

Dimitar Molerov
Dimitar Molerov is a research assistant at the Humboldt-University Berlin, Germany and pursuing his doctorate. He is currently working on a project focusing on the valid assessment of critical thinking (in online contexts) of students in higher education. In addition, he also has worked on the coordination project of the BMBF-funded research program "Modeling and Measuring Competencies in Higher Education (KoKoHs)".