# RESEARCH

# **Open Access**

# Developing and validating an online situational judgment test on the stress coping competence of nursing apprentices



Julia Warwas<sup>1\*</sup>, Wiebke Vorpahl<sup>1</sup>, Susan Seeber<sup>2</sup>, Philine Krebs<sup>2</sup>, Ulrike Weyland<sup>3</sup>, Eveline Wittmann<sup>4</sup>, Larissa Wilczek<sup>3</sup> and Aldin Strikovic<sup>4</sup>

\*Correspondence: julia.warwas@uni-hohenheim.de

 <sup>1</sup> University of Hohenheim, Fruwirthstr. 47, 70599 Stuttgart, Germany
 <sup>2</sup> University of Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany
 <sup>3</sup> Westfälische Wilhelms-Universität Münster, Georgskommende 26, 48143 Münster, Germany
 <sup>4</sup> TUM School of Social Sciences and Technology, Technical University of Munich (TUM), Arcisstraße 21, 80333 Munich, Germany

# Abstract

Converging evidence of high stress levels in the nursing profession as well as their detrimental effects on the execution of nursing actions, guality indicators of care, and self-reported job satisfaction justify the need for test environments that assess stress coping competences in dealing with domain-specific stress factors. Reliable and valid information on the test takers' stress coping competence can help to identify individual deficits and supportive needs and thus, to design training measures that may contribute to job retention and a skilled workforce. This article describes the development and validation of a digital assessment that ensures authenticity by means of video-based situational judgement tests. Analyses of the performance data were conducted in four steps: (1) the scaling of the data with a partial credit model in order to identify items that are subject to measurement error, (2) differential-item-function analyses for different specializations in nursing education, (3) dimensionality tests, and (4) correlations between test performance and perceived psychological work stress in order to gauge criterion validity. The total sample consisted of n = 300 German nursing apprentices. The results show that the developed test instrument can be used for a reliable and valid measurement of coping skills of (prospective) nurses, covering two essential dimensions. Limitations of the current study and related objectives of future research are discussed together with questions of practical applicability.

**Keywords:** Online situational judgment testing, Stress coping competence, Test validation, Nursing education

# Introduction

Even before the COVID-19 pandemic, health care professionals were considered to be (over-) loaded with specific stressors in their field of activity, including terminal care, resistance from patients and disputes with relatives, as well as restrictive working conditions (cf. Sarafis et al. 2016). The pandemic has aggravated these factors, raising concerns about the emotional and physical well-being of nurses even further (Schulze and Holmberg 2021). In general, occupational stress can lead to dissatisfaction, psychosomatic disorders, physical complaints, and burnout (Schuster et al. 2011; Krause and



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

Dormagen 2007). In the nursing profession, the consequences of stress become evident in growing numbers of overtime hours and sick days and an increased severity of illnesses among nursing staff (Isfort and Weidner 2018). Additionally, research has shown that stress reduces the level of compassion for patients and the quality of care due to more frequent mistakes (Sarafis et al. 2016). Therefore, coping effectively with stressors and their consequences is a valuable goal not only for nurses themselves but also for the health care system of a society.

The development of a theoretically sound, reliable, fair and valid instrument for testing stress coping competence contributes to attaining this goal in two ways: On the individ*ual* level, profound knowledge on the strengths and weaknesses that a prospective nurse possesses in dealing with the specific stressors of his/her future professional field helps to take preventive measures. At this early stage of professional development, these measures can be an integral part of social interaction and educational mediation in high-quality processes of Vocational Education and Training (see Böhn and Deutscher 2021). Put differently, these measures may consist in focused attention, needs-oriented guidance and support from instructors and team members at the workplace as opposed to remedial training measures in advanced stages of professional development and health complaints (Flaxman and Bond 2010; Westermann et al. 2014). On a system level, using the test instrument in large-scale assessments can enrich evaluations of curricular reforms aiming to enhance multifaceted vocational action competences or extensive programs aiming to build resilience among all professionals. From this perspective, stress coping competence represents a facet of personal competence or rather self competence, which plays a discrete role in skilled activities just like the facets of technical/expert, methodical, and social competence (e.g. Frey and Ruppert 2013; Kaspar et al. 2016).

While generic questionnaires exist that allow for self-assessing one's own stress management, there are currently no empirically approved stress-related competence tests tailored to nursing professionals. Available questionnaires do not inquire how respondents deal specifically (and potentially variedly) with a delineated range of contextualized, stressful situations of this activity domain. They rather focus on habitualized ways of handling feelings of stress in general or of handling a subjectively recurring or dominant stressor that a person pictures in her/his mind while filling out the questionnaire, such as social conflicts. Possible answers consequently offer generalized descriptions of prevailing thoughts, emotions and behavioral tendencies, such as self-blame, distraction, or seeking support (Kato 2015). With respect to the level of fidelity of stimuli and responses (i.e., authenticity), questionnaire-based diagnostic instruments of professional knowledge and abilities are clearly inferior to work samplings and, within technologybased test environments, to simulations of the task- and situation-specific demands of a particular workplace (e.g., Tuzinski 2013). However, both ethical and practical research issues impede simulations of (over)straining work situations in the nursing domain, which is a highly interactive, even intimate, service for and with human patients or clients (Glaser 2006; Kaspar et al. 2016; Bühle and Weihrich 2020). In the medical domain, situational judgment tests provide an established alternative, ensuring high levels of both authenticity and content validity (Patterson et al. 2016; Reiser et al. 2022). In the nursing domain, a few articles document situational judgment tests for core *functional* activities of nursing, such as conducting hygiene procedures (Heier et al. 2022; Heininger et al.

2021) or performing recipient-directed activities towards specific groups of patients/ clients (Kaspar et al. 2016). Yet, comparable approaches to measure self-directed competences, such as coping effectively with stressors and stress reactions that are inherent in or accompany nursing work, are missing. To fill this gap, we have dedicated a distinct part of the current research project EKGe<sup>1</sup> to modeling and measuring the stress coping competence of nursing apprentices by developing situational judgment tests that offer authentic video vignettes of stressful work situations in a digital test environment.

The present paper reports central steps in the construction and validation of this test instrument for measuring stress coping competence in nursing (CopeCo-N). In the logic of an *evidence centered design (ECD)* for test environments according to Mislevy (2013), this report concentrates on the domain analysis and domain modeling and on the conceptual assessment framework. Covering domain analysis and modeling in a first step, section "Modeling stress coping competence" introduces this complex construct of domain-specific demands. The theoretically established four-dimensional model then serves as the student model in the conceptual assessment framework. Section "Design of the stress coping competence instrument for nursing (CopeCo-N)" describes the framework by presenting approaches to and results from the task and evidence models. The task model refers to the design of tasks and items to elicit responses that demonstrate stress coping competence, while the evidence model explicates the scoring rules (cf. Mislevy 2013). Section "Methods" explains the methods underlying the validity checks for the CopeCo-N before section "Results" presents the obtained results. The database for examining the validity of the test instrument is a sample of 300 apprentices from three specializations in nursing education.

Using partial credit modeling, we are interested, firstly, in the quality of each developed test item. Therefore, we identified singular items that are prone to measurement error to maximize the reliability of the test instrument. Second, the CopeCo-N should be appropriate and fair for all three specializations in nursing education. For this reason, we use differential item functioning (DIF) analyses to examine whether item difficulties for these subgroups even out across the test and exclude items that have a strong subgroup difference and, to that extent, discriminate against a subgroup. Third, we are interested in whether the dimensionality assumed in the theoretical modeling of stress coping competence can be empirically represented with the data. The fourth objective is to test the criterion validity of the developed test through correlation analyses.

## Modeling stress coping competence

*Domain modeling* in the ECD framework (cf. Mislevy 2013), i.e., outlining the essential dimensions of a complex construct of competence in a designated area of professional practice, is preceded by *domain analysis* in order to specify the demands that should be mastered through deliberate, knowledgeable and skilled behavior in exactly this area (Shavelson 2012). In the project EKGe, structured group interviews with experts, following a Delphi method (de Meyrick 2003) in three rounds, provided the informational

<sup>&</sup>lt;sup>1</sup> Erweiterte Kompetenzmessung im Gesundheitsbereich [Extended competence assessment in the healthcare sector; a project from the research program ASCOT+; see https://www.ascot-vet.net/ascot/de/home/home\_node.html for detailed information].

ground of domain analysis. These interviews served to (1) assemble stressful situations in different nursing settings, (2) consensually choose a broad set of situations that any nurse likely encounters in the first years of professional practice and cannot judge and handle simply by adhering to standard *technical* rules of the profession, and (3) consensually define the constituents of each stressful situation in terms of the stress factors, participants, and organizational features of the related nursing action and its typical setting, as well as the most common psychosomatic stress reactions. The expert group consisted of 19 representatives of nursing practice, of nursing and educational sciences, of schools and of regulatory policy. This multitude of perspectives ensured high levels of typicality and authenticity of the stress situations for the focal domain. Furthermore, the guidelines for selecting and refining the stressful situations closely aligned with relevant taxonomies of work science. The latter concern action-regulatory conceptions of stressors as situationally embedded factors that impede, interrupt or interfere with an individual's execution of planned professional activities (Hacker 2020). These factors can be classified into stressors residing in task features (such as task incompleteness), working conditions (such as working overtime), social relations (such as conflicts) and working environment (such as a lack of work equipment) (Metz and Rothe 2017).

Despite compilations of stress trainings (Heinrichs et al. 2015) and integrative reviews of *Competence Based Education* in nursing (Lavoie et al. 2018), our literature search yielded no specific models of stress coping competence in the nursing profession that could be readily adopted for the purpose of domain modeling. Thus, Deutscher and Winther's (2022) synthesis of competence concepts in Vocational Education and Training provided a starting point. The authors conclude that different concepts converge in

"an understanding of vocational competence being a latent construct (or most often a conglomerate of several constructs) that allows people to act in various vocational situations in a vocational domain. E.g. Mulder et al. (2006, p. 82) define the concept of vocational competence as 'the capability to perform by using knowledge, skills, and attitudes that are integrated in the professional repertoire of the individual." (Deutscher and Winther 2022, p. 307).

The authors further point out that most concepts of competence presume (inter alia) that competence *enables* effective performance of vocational actions, and that it encompasses all levels cognitive *processing* in order to *transform* factual and rules-based knowledge into targeted action that is sensibly adapted to current conditions or demands. Thus, conceptual, procedural, and interpretational knowledge "represent an action schema for performing vocational tasks" (Deutscher and Winther 2022, 307), which is fueled by motivational, emotional and volitional facets of competence when translating the schema into factual operations. In a similar vein, Lavoie et al. (2018, 240) define competence in the nursing profession as "a complex knowing of how to act based on the effective mobilization and combination of a variety of internal and external resources in a family of situation". They refer to a *Competency Outcomes and Performance Assessment Model* as one feasible approach to delineate "core nursing competencies", which feature assessment and intervention, communication, critical thinking, teaching, human

caring relationships, management, leadership and knowledge integration (ibid.).<sup>2</sup> Finally, Kaspar et al.'s (2016) *model of geriatric nursing competence* spans recipient-directed, institutional/organizational and self-related competencies. However, the explicit focus of their contribution lies on the recipient-directed pillar, which is set out with detail and rigor (diagnosis and reflection, praxis and technique, interaction and communication). The authors also emphasize that the *self-directed* pillar deserves more attention in future studies aiming to elucidate the self-caring "health literacy" of nurses.

Against this background, our approach to model stress coping competence integrates more specific assumptions and concepts of scientifically evaluated stress trainings, stress theory, and coping effectiveness research. Prevailing transactional explanations of the emergence of stress reactions all deal with *coping strategies* that individuals use after appraising a situation as harmful, threatening or challenging (Lyon 2011). Thus, Lazarus and Folkman (1984) define coping as the constantly changing cognitive and behavioral efforts to handle (acute or persistent) external and/or internal demands that are subjectively assessed to strain or exceed individual resources. Nevertheless, the widespread use of questionnaire-based coping assessments (see section "Introduction") has fueled dichotomous conceptions of habitualized coping styles spanning all or at least similar stress situations (Jang et al. 2007; Kato 2015). Among those dichotomous conceptions is the distinction of (supposedly superior) problem-focused and (supposedly inferior) emo*tion-focused* coping styles, which seem to be mutually exclusive only when adopting the perspective of habitual tendencies. When dealing with situationally embedded stressors (i.e., during an ongoing stress episode), both strategies fulfill important complementary functions. Emotion-focused coping strategies, especially self-soothing, are often necessary to "cool down" in a first step and to release the mental and energetic resources needed for targeted and deliberate problem solving (Kaluza 2011). Moreover, emotionfocused strategies are indispensable to deal with stressors that cannot be changed, such as human loss, or when helpful resources are not available (Adams et al. 2011). Another long-established dichotomy of coping is the distinction between functional vs. dysfunctional strategies. In particular, strategies of resignation, escape, and substance abuse are classified as dysfunctional or maladaptive (Heinrichs et al. 2015). This distinction is valuable in that dysfunctional strategies have consistently proven to be harmful to wellbeing, especially in the long term. However, the remaining functional strategies deserve to be further subdivided according to which functions they fulfill specifically.

More recent, nonprescriptive classifications distinguish coping strategies with respect to the stress-inducing aspects of a *transactional* stress episode which they tackle primarily: *instrumental coping, mental coping* and *palliative-regenerative coping* (Heinrichs et al. 2015). In working contexts, *instrumental* coping aims at reducing or altering stressors, i.e., the mostly external causes that impede, interrupt or interfere with an individual's execution of planned professional activities (Hacker 2020). These causes may reside

<sup>&</sup>lt;sup>2</sup> Lavoie et al. (2018) further point to a framework of *Quality and Safety Education for Nurses' competencies* that lists patient-centred care, teamwork and collaboration, evidence-based practice, quality improvement, safety and informatics as core competencies. In light of the ECD framework for the construction and implementation of competence tests that structures the present article, these five components are more related to domain analysis than domain modeling. It should be noted that even the core competencies listed above would take an ambivalent place in the ECD framework since, for example, critical thinking is more consistent with notions of individual prerequisites or enablers of professional action than, for example, management.

in working tasks, working conditions, social relations, and/or the working environment (Metz and Rothe 2017). Furthermore, instrumental coping comprises activities that intend to mobilize or build up internal and/or external resources that help to carry out working activities successfully again, such as developing skills and effective routines or seeking help from colleagues. *Mental* coping accounts for the fact that each individual appraises situational demands potentially differently, such that subjective evaluations might render a particular stressor as merely irritating or rather as harmful, dramatic, etc. (Lazarus and Folkman 1984). Mental coping therefore aims to change one's perception and evaluation of a situation deliberately (i.e. purposeful reappraisal), for instance by using methods of cognitive restructuring. Experienced stress reactions such as fright or exhaustion, i.e. the immediate or subsequent psychosomatic consequences of stressors, are at the center of *palliative-regenerative* coping. The aim here is to reduce those reactions via relaxation methods and recovery, starting within the current stress episode (as in the case of self-soothing) or afterwards (as in the case of a meditational unit at home).

Another recent development in research on coping effectiveness that overcomes the traditional focus on cross-situational, habitualized, or preferred coping styles is the concept of coping flexibility. *Coping flexibility* refers to an individual's ability to effectively change and match his or her coping strategy according to the features of the stressful situation at hand (Kato 2012). A situationally appropriate strategic choice integrates short- and long-term options, types of stressors, and available resources in the current situation. Usually, a stressful situation is a multifaceted affair that requires a combination of different types of coping strategies. Thus, effective coping can be characterized by the broadest possible repertoire of available strategies, and regarding their implementation, by a mindful balance among instrumental, mental and palliative-regenerative coping strategies, plus the flexibility to alter strategies as situational demands change (Kaluza and Chevalier 2016).

Consequently, stress coping competence as opposed to habitualized, preferred coping styles includes both differentiated situational perceptions (i.e. the recognition of current stress-inducing causes and their potential psychosomatic impact) and flexible strategic choices that are tailored to the demands and resources of the current situation, and combined to reduce stressors and regulate stress reactions (i.e. the 'designing' of an appropriate, often multifaceted approach of coping in that particular situation). Put differently, stress coping competence denotes the ability to understand a stressful situation in the light of its constitutive features and to select and purposefully implement coping strategies that together are adequate for the situation and supportive of the individual ('Erkennen und Gestalten', see Joiko et al. 2010). A person must be able to assess, for example, whether a stressor is modifiable (such as chaotic working processes in a nursing team) or unalterable (such as the death of an incurable patient), whether and which external resources (such as colleagues) are available, etc. On this basis, a person can choose appropriate instrumental, mental and/or palliative-regenerative coping strategies. Drawing on a broad knowledge base of these coping strategies, this selection usually combines short-term, immediate strategies (such as positive self-instruction) within the situation and long-term strategies subsequent to the situation for recovery (such as fulfilling leisure activities) and prevention (such as restructuring working processes in the team). With a claim of *professional judgment* on authentic situational vignettes

(Taylor 2006), an individual's ability to *justify* his or her strategic choices can be considered an additional aspect of reasoned and targeted rather than habitualized coping and has the potential to enhance flexible ways of coping in the long run.

In essence, stress coping competence comprises individual knowledge and skills for dealing effectively with psychological stressors that typically occur in situations of professional nursing practice as well as for regulating the resulting psychological stress reactions, both with the aim of maintaining one's own productive capacity, well-being, and health. Our modeling approach (see above) is compatible with notions of *Competence as Situated Professionalism*, implying that the focal competence draws its meaning from a certain context of professional practice and that the given context provides the possibilities ('affordances') and constraints for developing and for demonstrating competent behaviors (Mulder 2014). We distinguish four competence dimensions as the basis for our student model in the ECD framework (cf. Mislevy 2013), which will be examined by means of confirmatory factor analysis (CFA) subsequently:

- 1. Recognition of situation-bound stressors when planning or conducting nursing activities
- 2. Recognition of stress reactions that likely accompany or follow this situation
- 3. Choice of (short- and long-term) coping strategies that are adequate to diminish the relevant stressors and regulate stress reactions
- 4. Justification of the chosen coping strategies.

 and (2) refer to an appropriate assessment of the stressful situation in nursing work and its consequences for one's own performance capacity, well-being and health. (3) and
 relate to the appropriate selection and justification of ways to reduce or prevent detrimental effects on performance capacity, well-being and health.

## Design of the stress coping competence instrument for nursing (CopeCo-N)

Technology-based methods are particularly suitable for measuring professional competencies because they allow working situations to be presented in an authentic but standardized manner, which is important for covering large samples. At the same time, they can overcome the disadvantages of traditional testing methods (e.g., paper–pencil tests), such as a lack of professional relevance or insufficient situational awareness. Moreover, they ease the mapping of important facets of competence, namely, the interpretation of situations and the extraction of information from situations for professional decisions and actions. In line with Shavelson (2012), a competence measurement that allows robust statements about coping with requirements in the "real" or professional world should meet the following standards:

"A measure of competence should tap complex physical and/or intellectual abilities and skills to produce observable performance on a common standardized set of tasks that simulate with high fidelity the performances that are expected to be enacted in the «real world» («criterion») situations to which inferences of competence are to be drawn, with scores reflecting the level of performance (mastery or continuous) on tasks where improvement can be made through dispositions for selfregulation, learning, and deliberative practice." (Shavelson 2012, 78) Consequently, the measurement of professional competencies necessitates that the occupational field of action and its requirements are determined as precisely as possible and that the measurement procedures are oriented "to the respective work activities and processes" (Klotz 2015, 18) when constructing a set of tasks to be solved by (aspiring) professionals in this field. Thus, moving to the next steps of the ECD framework, we present the design of the CopeCo-N in terms of the task model and evidence model, based on the student model shown in section "Modeling stress coping competence" (cf. Mislevy 2013).

With the aim of valid task modeling, we conducted interviews again with experts from nursing science and education, care institutions and regulatory policy to specify authentic and potentially stressful care situations in greater and lively detail. We also made sure that the situational features corresponded with criteria for analyzing situation-bound nursing activities used in nursing education (Hundenborn 2007), covering nursing occasion/cause, interaction structures, institution, and individual experiences and conduct in the focal situation. Considering these characteristics, we developed scripts for video vignettes that represent stressful situations and culminate in a moment in which the protagonist has to make decisions and take actions. To ensure authenticity, experts from science and practice inspected and commented the scripts in another round of review. This round included testing whether the stressful stimuli were perceived by viewers from the professional domain as intended by the research team. The feedback led to modest changes before shooting the videos with professional actors.

In addition, we conducted comprehensive curricular analyses to establish the content validity of the test instrument. The entire range of nonacademic nursing education in Germany, spanning geriatric nursing (Altenpfleger\*in), clinical nursing (Gesundheitsund Krankenpfleger\*in) and nursing specialist (Pflegefachmann\*frau), was considered in this step. Eventually, we developed and filmed nine vignettes of stressful situations residing in three fields of practice of care: (1) nursing homes for elderly individuals, (2) hospital care and (3) ambulant care.<sup>3</sup> The following table shows the stressful situations included in the final test instrument, sorted by the respective fields of practice (Table 1).

On average, the completed video vignettes<sup>4</sup> depicting stressful situations in the nursing profession have a length of 2 min. To enhance authenticity and immersion into the three institutional fields of care even further, three short introduction videos were created in addition. These introduction videos contain vivid information about the working contexts (such as colleagues) and thus, facilitate the recognition of relevant constituents of a situation in the following test stimuli.

To translate our understanding of stress coping competence into a psychometric model, we developed nursing-specific *items for the situational judgement test* that

<sup>&</sup>lt;sup>3</sup> The nine situations are relevant to all three institutional fields of nursing practice. However, in order to illustrate the situations in an authentic setting, they were assigned to particular fields of practice as prototypical examples of severe and recurring stressors in these settings on the basis of the group interviews with experts. Being confronted with the process of dying, for instance, can occur both in hospitals and nursing homes for elderly individuals. Nevertheless, the death of clients/patients sadly is a regular occurrence in nursing homes (but comparably exceptional in hospitals) and typically aggravated as a psychological stressor by the fact that nursing activities for the elderly include the build-up of trustful interpersonal relations over several months or years. Note that the generalist nursing training in Germany now gives prospective nurses the opportunity to provide professional nursing care in all three fields.

<sup>&</sup>lt;sup>4</sup> All video vignettes are available and classified according to the depicted constituents of stressful situations of nursing as an *open educational resource* (https://ekge.de/).

Table 1         Overview of stressful situations in the final test instrument sorted by fields of plant	oractice (FP)
---	---------------

FP 1: Nursing homes for elderly individuals

Sit. 1.1: Being confronted with dying

Sit. 1.2: Setting priorities in areas of conflict between ethical issues

Sit. 1.3: Role-based relationship management in dealing with colleagues

FP 2: Hospital care

Sit. 2.1: Setting priorities in the face of parallel care processes

Sit. 2.2: Being confronted with suffering

Sit. 2.3: Developing professional identity in dealing with tensions between professional and private expectations

FP 3: Ambulant care

Sit. 3.1: Role-based relationship management in dealing with relatives

Sit. 3.2: Dealing with work overload

Sit. 3.3: Dealing with errors in nursing activities

measure the competence levels of apprentices. For each stressful situation, we developed several items that relate to the theoretically derived four dimensions:

- 1. the recognition of stressors ("What are the reasons why this situation is stressful for > name of nurse in the video < ?"),
- the recognition of stress reactions ("How does>name of nurse in the video < experience this situation?"),
- 3. the choice of coping strategies ("Which measure(s) do you think are most appropriate <u>now</u> to deal with the demand(s) of this situation?"; "Which measure(s) do you think are particularly appropriate for dealing with such demand(s) in the <u>long term</u>?"),
- 4. the justification of coping strategies (including arguments about the stressor-reducing and/or stress-regulating effects of advisable strategies: ">Name<should ask his colleague>Name<for help so he can check the needs of the other residents and stop feeling overwhelmed by concurrent demands.").</p>

The items reflect the dimensionality assumed in the competence model and are thus consistently aligned with it. With regard to test economy and a standardized, automated scoring of answers in a large-scale implementation of the test, the items have a multiple response format. Response options capture a wide range of nursing-specific stressors, stress reactions, and coping strategies. For each stressful situation, the situational judgement test includes three to six items that correspond with the delineated competence dimensions but are delivered in a simplified language. A total of 9 items measure the recognition of situation-bound stressors, 9 items cover the recognition of stress reactions, and another 5 items aim to assess their justification. Table 2 shows an exemplary item for a situation in a nursing home for elderly individuals pertaining to the first competence dimension.

The *scoring* of possible responses links the student and task models, thus establishing the *evidence model* via the CFA. In our study, we used multi-stage scoring. If a nursing apprentice did not select a correct answer or selected all answering options, we awarded zero points. We counted one point if a test taker selected one of two correct answers. Two points were attained if all correct answers were marked as such. The CopeCo-N

 Table 2
 Exemplary item for the recognition of stressors for Situation 1

What are the reasons why the si	tuation is stressful for Justus?
---------------------------------	----------------------------------

A. Justus is alone and gets no support

B. Justus has too little time to take care for Mrs. Reiter adequately

C. Justus neglects the care for Mrs. Reiter

D. Justus neglects the care of other residents

E. Justus is making a mistake by leaving Mrs. Reiter alone right now

F. Justus is personally confronted with the fact that the relationship between him and Mrs. Reiter is about to end

	N I I	C • •	·						
Isble 2	Numbor	of cituations	itome and	mayimiim	noint	COROC CORTO	d hv	nurring	cottinac
I a DIE S									
		01 5100001157	1001110 0110		0000000	500105 50100	$\sim$ ,		50000
		,							

Fields of practice (FP)	Number of situations	Number of items	Maximum point scores
FP 1: Nursing homes for elderly indi- viduals	3	12	24
FP 2: Hospital care	3	10	20
FP 3: Ambulant care	3	10	19
Total	9	32	63

consists of 31 polytomous tasks and one dichotomous task. In total, a maximum of 63 points can be achieved. Table 3 gives an overview of the number of situations, items and maximum achievable scores per field of practice.

## Methods

#### **Test implementation**

In light of ethical and pragmatic issues of researching nursing activities set out in section "Introduction", a video-based situational judgment test (SJT; e.g., McDaniel and Nguyen 2001; Chan and Schmitt 2005) seemed most appropriate to assess stress coping competence of (prospective) nursing professionals. In general, the advantages of the SJT are that the construct of interest is measured in situations that are critical for a successful outcome. Relevant response options can be formulated to capture conceptual, procedural, and interpretational knowledge of correct or desirable behavior. We used "should-do instructions" when asking participants to select and evaluate effective coping strategies in the respective situations. We then determined varying qualities of responses in terms of appropriateness and effectiveness through an intensive review of a range of answer alternatives (i.e., differing appraisals of situation-specific stressors and potential stress reactions; differing combinations of strategic combinations and their respective justifications), conducted again by domain experts from nursing science and nursing education and aligned with the theoretical basis of effective and flexible coping (section "Modeling stress coping competence"). Thus, the SJT we implemented has the strength of representing complex stressful situations in the nursing profession in great detail, authentically, efficiently and economically. Yet a drawback is that the subjects' responses might reflect behavioral intentions more than actual (or simulated) behavior.

Specializations in	Gender		Age group			Country of birth	Total
nursing education	f	m	<21	21–25	>25	Germany	
Geriatric nursing	79.0	20.3	21.0	37.0	42.0	81.2	138
Clinical nursing	78.0	18.3	48.8	32.9	18.3	93.9	82
Nursing specialist	85.2	11.1	27.8	44.4	27.8	87.0	54
Total	79.9	17.9	30.7	37.2	32.1	86.1	274 <sup>a</sup>

**Table 4** Specializations in nursing education by gender, age group and country of birth (in percentage)

<sup>a</sup> For n = 26 students, data on gender, age, and country of birth are missing

Note. Nursing specialist training began approximately six to eight months after the start of training in geriatric nursing and clinical nursing

The test was embedded in an open-source learning content management system that the nursing apprentices could access online (ILIAS). They watched the videos and completed the tasks on tablets using headphones. In this way, the nursing apprentices were able to proceed at their own pace. Test implementation was technically prepared and supervised. For this purpose, test booklets were assembled to vary the query sequence.

#### Sample

To validate the test instrument for assessing stress coping competence of nursing apprentices, we used data from the first measurement point of a longitudinal study of the research project EKGe. Data collection in a test booklet design took place between April and July 2021 in North Rhine-Westphalia (Germany). A total of 300 nursing apprentices in the educational specializations of "geriatric nursing" (Altenpfleger\*in) (n=152), "clinical nursing" (Gesundheits- und Krankenpfleger\*in) (n=83) and "nursing specialist" (Pflegefachmann\*frau) (n=65), who were in their second year of training at the time of measurement, participated in the test. Female trainees were prevalent in the sample (79.9%). The age of the respondents was distributed heterogeneously. Nearly 14% of the trainees was born in a country other than Germany. Table 4 provides an overview of the sample.

Overall, the respondents achieved an average of 21.5 points in the test out of a maximum of 63.0 possible points. A comparison of test results among the different specializations in nursing education indicates a better performance of students in clinical nursing. While they achieved an average of 24.8 points in the test, the geriatric nursing students and nursing specialists attained 20.5 and 19.9 points, respectively. However, the comparably lower test performance of the specialist-nursing students could be due to a shorter training period of this group at the time of the testing. Although all students in the sample were in their second year of occupational training, the training for nursing specialists began approximately 6 months later than training for geriatric nursing and clinical nursing.

#### Data analysis

The analysis of the performance data for the measurement of stress coping competence was conducted in four steps. In the *first* step, the performance data were scaled using a one-dimensional partial credit model to examine the quality of the test items through

common and probabilistic fit indices, in particular, weighted mean squared (wMNSQ), t statistics, and corrected item-total correlation (discrimination). Following the recommendations of Gnambs and Nusser (2019), we consider items with a wMNSQ between 0.8 and 1.2 as acceptable. Additionally, we consider t values greater than |2| as less compatible with the model than expected (p<0.05) (Bond et al. 2021). However, because the t statistic is highly dependent on sample size (see Smith et al. 2008), we place more emphasis on the wMNSQ. In addition, we used item-total correlations. Following common rules of thumb for assessing correlations of item scores with the total score, values greater than 0.2 are considered acceptable (Pohl and Carstensen 2012). To maximize the reliability of the test instrument, items with measurement errors were excluded iteratively. All analyses were performed using Conquest 2.0 and SPSS 27.

*In the second step*, we conducted *DIF* analyses to ensure that the test items were appropriate for different specializations in nursing education. To assess subgroup invariances, we refer to the NEPS study's general recommendations for assessing DIF. According to Pohl and Carstensen (2012), we consider absolute differences in estimated difficulties greater than 1 logit to be very strong DIF and absolute differences between 0.6 and 1 to be worthy of attention for further investigation. To ensure that the CopeCo-N is fair for all subgroups of specializations in nursing education, items with strong DIF were excluded.

*Third*, based on these results for the exclusion of items with measurement errors, model comparisons were run to examine whether the theoretically assumed multidimensional structure of stress coping competence could be empirically represented with the data. Our four-dimensional model, elaborated in section "Modeling stress coping competence", classifies the items of the test instrument into four dimensions: (1) recognition of stressors, (2) recognition of stress reactions, (3) choice of coping strategies appropriate to the situation, and (4) justification of coping strategies. We assume the first two dimensions and the last two dimensions to be more interrelated, since they pertain to either an appropriate situational assessment or to appropriate ways of dealing with inherent demands and consequences of a situation. Dimensionality is examined by comparing the fit indices of a one-dimensional model (see Fig. 1, right panel). The deviance is used to evaluate model fit.

To examine criterion validity in the *fourth* step, we calculated correlations between the test performance of nursing students (modeled as WLE person parameters) and their subjective perceptions of psychological work stress. Experiences of stress were assessed in a survey via the two scales *emotional exhaustion* and *aversion* towards clients (Hacker and Reinhold 1999). Both scales consist of three items each and answers are given on a five-point Likert scale ranging from 1 = "disagree entirely to 5 = "agree very much". The reliability parameters (Cronbach's alpha) in the present dataset reach  $\alpha = 0.87$  for the *emotional exhaustion* scale.

## Results

#### One-dimensional partial credit model for identifying items subject to measurement error

Overall, the 32-item test instrument showed good EAP/PV reliability (0.86). The developed items cover a wide range of difficulty levels. The wMNSQ was within an



Fig. 1 One- and four-dimensional models of stress coping competence

acceptable range for all items  $(0.87 \le \text{wMNSQ} \le 1.16)$  (see Appendix Table 10). Nevertheless, the parameter values of two items lay in a problematic range. One item had unsatisfactory discriminatory power (discrimination = 0.13) (Item BK1211). After exclusion of this item, the t value of another item fell into a problematic range at 3.0 (Item BK14242). This item also was excluded.

Consequently, further analyses refer to 30 items. The EAP/PV reliability for the remaining item pool improved and reached a value of 0.891. The remaining items showed a very good fit ( $0.84 \le \text{wMNSQ} \le 1.13$ ), largely acceptable t values ( $-2.4 \le \text{t}$  values  $\le 2$ ), and satisfactory discriminatory power (discrimination  $\ge 0.26$ ) (see Appendix Table 11). For two items, the t values lie just above the threshold for item exclusion. However, since the model violation is not considered meaningful, these items remained. We report a renewed check of model violation later in this paper.

#### DIF analyses of the specializations in nursing education

Subgroup invariance analysis was conducted at both the global (i.e. cross-item) level and the local item level. The results at the global level yield no significant DIF between specializations in nursing education (see Appendix Table 12). However, it is apparent that students in clinical nursing scored slightly higher than their counterparts in the nursing specialist and geriatric nursing groups. Examination of the DIF parameters at item level revealed a mixed picture with alternating advantages between the three specializations in nursing education (see Appendix Table 13). One item had a strong DIF (Item BK1212), indicating a higher item difficulty for the geriatric nursing students. To ensure that the CopeCo-N was fair for all nursing subgroups, we excluded this item. After that, the EAP/PV reliability reaches a very good level (EAP/PV Rel. = 0.91). In addition, all item characteristics covered an acceptable range  $(0.87 \le wMNSQ \le 1.14; -1.8 \le t \text{ values} \le 2.1; \text{ discriminatory power} \ge 0.26).$ Table 5 presents an overview of excluded and remaining items according to institutional fields of nursing and stressful situations. The established institutional fields of practice and situations (following Wittmann et al. 2022) continue to be represented evenly by the remaining items.

Fields of practice (FP) and situations (Sit)	Number of items before exclusion	Number of items excluded	Number of items after exclusion
FP 1: Nursing homes for elderly individuals	12	0	12
Sit. 1.1: Being confronted with dying	6	0	6
Sit. 1.2: Setting priorities areas of conflict between ethical issues	3	0	3
Sit. 1.3: Role-based relationship management in deal- ing with colleagues	3	0	3
FP 2: Hospital care	10	1	9
Sit. 2.1: Setting priorities in the face of parallel care processes	3	0	3
Sit. 2.2: Being confronted with suffering	3	0	3
Sit. 2.3: Developing professional identity in dealing with tensions between professional and private expectations	4	1	3
FP 3: Ambulant care	10	2	8
Sit. 3.1: Role-based relationship management in deal- ing with relatives	3	2	1
Sit. 3.2: Dealing with work overload	4	0	4
Sit. 3.3: Dealing with errors in nursing activities	3	0	3
Total	32	3	29

 Table 5
 Distribution of items by fields of practice and situations (after item exclusion)

**Table 6** Number of items, intercorrelations and reliabilities for the subdimensions of stress coping competence (4-dimensional)

Subdimensions	1	2	3	Number of items	Reliability (EAP/PV)
1. Recognition of stressors	1			8	0.71
2. Recognition of stress reactions	0.87	1		7	0.71
3. Choice of occupation-specific coping strategies appropriate to the situation	0.82	0.82	1	9	0.79
4. Justification of occupation-specific coping strategies	0.72	0.72	0.87	5	0.72

#### **Dimensionality test**

Subsequently, model comparisons served to examine whether the theoretically assumed multidimensional structure could be empirically reproduced. Table 6 provides information on the number of items in each subdimension, the subtest reliabilities, and the intercorrelations when specifying the four-dimensional model. The subtest reliabilities of the four dimensions are in an acceptable range. While the dimension "justification of coping strategies" shows little intercorrelation with the dimensions of recognizing stressors and stress reactions, indicating relative independence, the two dimensions for choosing and justifying coping strategies correlate comparatively high with each other. Furthermore, the two dimensions of recognizing stressors and stress reactions, which runs counter to treating these content areas as separate test components.

Due to the theoretically assumed and empirically confirmed correlations between the first two dimensions of stress coping competence, referring to the assessment of stressful

Subdimensions	1	2	Number of items	Reliability (EAP/PV)
1. Recognition of stressors and stress reactions	1		15	0.78
2. Choice and justification of occupation-specific cop- ing strategies appropriate to the situation	0.83	1	14	0.90

**Table 7** Number of items, intercorrelations and reliabilities for the subdimensions of stress coping competence (2-dimensional)

## Table 8 Fit indices of dimensional analyses to the CopeCo-N

	1-dimensional	2-dimensional	4-dimensional
Reliability (EAP/PV)	0.91	0.78/0.90	0.71/0.0.71/0.79/0.72
Deviance	15378.279	15356.921	15361.264
Estimated number of parameters	59	61	68
Difference (to the 1-dimensional model)		21.35726 at 2 df; p < 0.001	17.01482 at 9 df; p<0.05

situations, and the last two dimensions, focusing on strategy selection and justification, it seems reasonable to examine a two-dimensional model additionally. The first dimension of this model includes 15 items from the realm of recognizing stressors and stress reactions. The second dimension includes 14 items in terms of stress management. Table 7 displays subtest reliabilities and intercorrelations in this two-dimensional model. While the first, appraisal-related dimension reaches acceptable reliability, the second, strategy-related dimension even attains a very good value. Although the level of intercorrelation (0.83) between the two subdimensions alone would also justify a one-dimensional model for measuring stress-coping competence, it seems tolerable in order to provide a more differentiated view and thus, additional insights into the test takers' existing or deficient capabilities with improved reliability over the four-dimensional model.

The results of model comparisons are shown in Table 8. The fit indices corroborate that the theoretically assumed multidimensional models are both superior to the one-dimensional model. The differences between the one-dimensional and the multidimensional models are significantly in favor of the multidimensional approaches. Although the difference between the two- and four-dimensional models is not significant, both the fit indices and the limitations presented for the four-dimensional model tend to support the superiority of the two-dimensional approach.

#### **Criterion validity test**

To examine criterion validity, we calculated correlations between performance in the test instrument for stress coping competence, modeled by the WLE person ability parameters, and the *emotional exhaustion* and *aversion* scales (see Table 9). When scaled unidimensionally, the WLE person ability parameters ranged from -4.46 to 1.03 with a mean of -0.60 (SE=0.88). For the two-dimensional scaling, the WLE person parameters

1	2	3	4	5	м	SE
1					- 0.60	0.88
0.84**	1				- 0.83	0.85
0.94**	0.64**	1			-0.34	0.93
- 0.15*	-0.15*	-0.16*	1		2.53	1.09
-0.25*	-0.20**	-0.27**	0.57**	1	1.77	0.87
	<b>1</b> 0.84** 0.94** - 0.15* - 0.25*	<b>1 2</b> 1 0.84** 1 0.94** 0.64** - 0.15* - 0.15* - 0.25* - 0.20**	1         2         3           1	1         2         3         4           1	1         2         3         4         5           1	1         2         3         4         5         M           1         -0.60         -0.84**         -0.83         -0.94**         -0.64**         -0.34           -0.94**         0.64**         1         -0.34         -0.34           -0.25*         -0.20**         -0.27**         0.57**         1         1.77

 Table 9
 Pearson
 correlations
 between
 performance
 on
 the
 CopeCo-N
 (1-dimensional, 2-dimensional) and the experienced stress scales

<sup>a</sup> Performance on the test instrument measuring stress coping competence was modeled using WLE person ability parameters. Significant correlations are marked with \*p < 0.05, \*\*p < 0.01; SE: Standard error; M: Mean

ranged from -3.09 to 0.80 with a mean of -0.83 (SE = 0.85) for the first dimension and from -4.02 to 1.65 with a mean of -0.34 (SE = 0.93) for the second dimension. Performance in the stress coping competence test (WLE person parameters) correlates significantly negatively, though weakly, with the experienced stress scales. Thus, higher person ability on the CopeCo-N is associated with lower perceived *emotional exhaustion at work* and lower feelings of *aversion towards persons requiring care*. Differentiation into two separate test components reveals that the second, strategy-related dimension correlates more strongly with the scales on experienced stress. In particular, the aversion scale shows a near-medium negative correlation with the WLE person parameters of the second dimension of stress coping competence. Thus, the theoretically hypothesized significant negative relationships between (prospective) nurses' stress coping competence as measured by the developed SJT and their levels of emotional exhaustion and aversion to patients/clients as reported via self-assessment are evident. Possible reasons for these systematic yet modest relations will be discussed at the end of the paper.

## Discussion

Despite increasingly complex demands in the nursing profession, which is characterized by specific stressors, a standardized yet authentic test instrument for assessing coping skills for these domain-specific stressors is missing. The aim of this paper was to close this gap. We developed the *video-based situational judgement test* CopeCo-N and examined its potential to deliver valid, differentiated and fair diagnostic information on the stress coping competence of nursing apprentices.

## Psychometric quality of the instrument

Using partial credit modeling and DIF analyses for the existing specializations in nursing education at the time of measurement, we were able to identify a pool of 29 items that have satisfactory item characteristics, do not discriminate one educational program against another, contribute to reliability, and ensure content validity. The test items further have a very good internal consistency. With an item processing time of approximately 40 min, the video-based situational judgement test CopeCo-N is also satisfactory from the perspective of test economy, considering its higher fidelity for professional activities compared to questionnaire-based (self-)assessments. An analysis of dimensionality shows that the theoretically substantiated multidimensional models of stress coping competence are superior to a one-dimensional model. Furthermore, the two-dimensional model achieves a better fit with the empirical data than the four-dimensional model. This two-dimensional model, covering (1) *the recognition of stressors and stress reactions residing in or following from a working situation* and (2) *the choice and justification of coping strategies* also aligns well with conceptual distinctions in the literature on stress, strain, and coping. In addition, significant negative correlations between obtained test scores and the test takers' self-reported feelings of emotional exhaustion and aversion towards persons requiring care support the criterion-related validity of the CopeCo-N.

The fact that correlations are systematic but only small to moderate in strength can have various reasons. First, methodological factors can reduce correlation parameters. Whereas the measurement of stress coping competence is based on performance scores in a situational judgment test, perceptions of work-related stress were recorded via selfassessment scales. Second, the constructs under study also differ from each other in their level of generality. Self-reported feelings of emotional exhaustion and aversion at work capture a small selection of symptoms for psychosomatic strain. Although these symptoms are prominent in the professional field of nursing, they still represent comparably generalized constructs since they result from various accumulated, stressful and overstraining work situations that haven't been mastered effortlessly. A prospective nurse's stress coping competence, in contrast, relates to the adequate recognition and mastery of several domain-specific and even situation-bound stressors as well as the regulation of related stress reactions. Last but not least, while the CopeCo-N concentrates on understanding and handling prototypical stressors and stress reactions in the main institutional fields of nursing (hospital care, ambulant care, nursing homes for the elderly), data collection for the present study took place under extraordinary surrounding conditions, namely the breakout of the COVID-19 pandemic. The latter may well have affected the participants' general levels of perceived exhaustion and aversion even though they might have demonstrated high abilities to understand and handle the prototypical stressors at their workplaces.

## Applicability and limitations

As outlined before, our measurement approach is based on the evidence centered assessment design (ECD; Mislevy 2013) with the starting point of domain analysis, which includes specific knowledge representations, skills, and abilities as well as the elaboration of typical tools, instruments, and interaction patterns in demanding situations that become relevant for test processing. However, for socially interactive work, as is the case of nursing, there are also technical and ethical limits to *simulating* situated professional actions, even more so in the inducement of stress and in the measurement of long-term coping skills. Consequently, although situations with different types of stressors can be represented well in video stimuli, the test subject remains outside the situation, so that the psychosomatic impact of stressful situations can be only elicited through the expressions and interactions of the protagonists. We dealt with this challenge by engaging actors with long-standing experience in the nursing profession.

The format of a video-based situational judgment test thus has the great advantages of authenticity and action orientation, but its structure implies the risk of violating the model requirement for local stochastic independence of the different items. Clearly, the items cannot be interpreted and answered decoupled from of their respective anchor situation. However, we checked the stochastic independence of the test items several times during item development and validation, finding no visual anomalies. Modelling an additional factor to cancel out the variance stemming from the item grouping in situations therefore did not seem compelling. When choosing SJT as the test format, its fidelity and coherence regarding real working situations and demands in the nursing profession received highest priority. Although decontextualized tasks might pave the way for even greater empirical differentiation between the proposed dimensions of stress coping competence (four dimensions versus two dimensions) they are almost incommensurable with conceptions of Situated Professionalism as outlined in the theoretical foundations of the competence model. An alternative approach for constructing the video-based SJT would be for each item to refer to only one stressful situation. With respect to test development (concerning the number of video stimuli), test economy, and the demands on the test takers' attentiveness to complex stimuli changing in high freguency, however, we have refrained from this option and limited the number of videobased, stressful situations.

Regarding the response format in the SJT, we opted for closed items with predefined response alternatives for several reasons. Undoubtedly, open answers would give even deeper insights into the participants' abilities to judge and handle stressful situations of nursing work and, thus, further enhance authenticity of the test format. On the other hand, the elaborate process of analyzing and evaluating given answers would restrict test economy, usage in big samples, and objectivity of results. It should be noted, however, that corresponding to the theoretical foundation of the test instrument, responses for strategy selection and justification are complex enough to reduce chances of guessing. As set out in section "Modeling stress coping competence", promising ways to tackle stressors and regulate stress reactions include a combination of coping strategies, so alternative response options reflected different combinations. Another argumentbesides test economy-played an important role in the decision to implement closed response formats: With open responses, written language skills are a test bias that should be controlled additionally. This influencing factor is particularly relevant for occupations in which trainees have a heterogeneous social, migration, and educational background. This is the case for the nursing professions.

A clearer limitation is that the sample comes exclusively from one federal state in Germany. Although trainees from Bavaria were included for a pretest of the developed instrument, the main survey took place only in North Rhine-Westphalia. And even though the curriculum for the newest educational program of the nursing specialist is valid throughout Germany, there may well be slight differences in the implementation of the older programs (geriatric nursing and clinical nursing) among the various federal states.

As a promising research perspective, the use of the test instrument in longitudinal designs should be further explored, e.g. to investigate the development of stress coping competence of (prospective) nurses over the course of their occupational training. In addition, these studies should examine robust predictors of competence gains, such as features of pedagogical interaction and support from teachers/trainers and members of nursing teams.

The practical value of the developed test instrument lies primarily in its reliable and valid diagnostic information on individual levels of stress coping competence. This includes information on particular strengths and deficits of prospective nurses, which could be addressed in early stages of their professional career, thus contributing to a preventive approach against training dropout or even burnout in later career stages. Given that the stressful situations depicted in the video stimuli represent stressors and settings that are typical for the work of nurses in other countries as well, the test instrument might also be used for comparative studies on the competence levels of entire groups of trainees in different systems of nursing education. Of course, such comparisons necessitate diligent domain analyses (see section "Modeling stress coping competence") in other countries in advance. Last but not least, the developed video stimuli can be integrated in instructional designs of nursing education, which often follow a situated approach to teaching and learning (Hundenborn 2007). Therefore, video stimuli will be made available as Open Educational Resources (https://ekge.de/).

## Conclusion

Although effective coping with stress in the nursing profession is of high relevance, standardized tests for assessing stress coping competence of (prospective) nurses are largely missing. In the present paper, we fill this research gap by using the video-based situational judgement test CopeCo-N. The CopeCo-N can help to identify potential deficits (and extant strengths) of apprentices regarding an important self-related facet of vocational action competence and to support them according to their needs.

#### Appendix

See Tables 10, 11, 12, 13.

ltem	Estimate	Standard error	wMNSQ	t-value	Discrimination
BK1111	- 0.086	0.053	0.98	- 0.2	0.43
BK1112	0.652	0.057	1.10	1.1	0.28
BK1113	- 0.966	0.053	1.11	1.8	0.37
BK1114	-0.117	0.052	1.01	0.2	0.45
BK1115	- 0.086	0.053	1.03	0.5	0.42
BK1116	-0.249	0.051	1.05	0.9	0.43
BK1131	0.077	0.054	1.00	0.1	0.39
BK1132	- 0.362	0.052	1.04	0.7	0.47
BK1133	- 0.458	0.062	0.95	-0.4	0.38
BK1141	0.068	0.053	1.16	2.4	0.28
BK1142	- 0.099	0.051	1.13	2.1	0.33
BK1143	- 0.985	0.055	0.99	- 0.2	0.43
BK1211	1.239	0.062	1.12	0.7	0.13
BK1212	0.055	0.062	0.92	- 1.8	0.43
BK1213	- 0.575	0.057	1.05	0.7	0.33
BK1221	0.528	0.056	1.11	1.4	0.27
BK1222	0.576	0.056	1.07	0.7	0.31
BK1223	- 0.273	0.051	0.99	- 0.2	0.47
BK1225	-0.186	0.052	1.10	1.7	0.38
BK1231	0.240	0.053	1.02	0.3	0.39
BK1232	0.147	0.053	0.95	- 0.7	0.48
BK1233	0.415	0.054	1.04	0.5	0.36
BK1341	0.275	0.054	1.01	0.2	0.42
BK1342	0.221	0.053	0.93	- 0.9	0.49
BK1343	-0.115	0.052	1.06	1.0	0.40
BK1351	0.727	0.057	0.90	- 1.0	0.48
BK1352	0.018	0.052	0.95	- 0.8	0.46
BK1353	0.072	0.053	1.00	0.0	0.49
BK14241	0.230	0.052	0.88	- 1.7	0.57
BK14242	-0.447	0.049	0.87	<b>- 2.6</b>	0.64
BK14243	- 0.385	0.050	0.97	- 0.5	0.50
BK14244	- 0.153 <sup>a</sup>	0.301	0.93	- 1.2	0.55

Table 10	Item information	for one-dimensional	partial-credit-model

Separation Reliability = 0.987; Chi-square test of parameter equality = 2127.33, df = 31, Sig. = 0.000

Problematic values are printed in bold

<sup>a</sup> The item parameters were fixed to zero as part of the item analysis. One parameter is fixed by ConQuest by default for model identification purposes

ltem Estimate		Standard error	wMNSQ	t-value	Discrimination		
BK1111	- 0.060	0.053	0.97	- 0.4	0.43		
BK1112	0.676	0.057	1.07	0.8	0.27		
BK1113	- 0.937	0.053	1.10 1.5		0.37		
BK1114	-0.091	0.052	0.99	- 0.1	0.46		
BK1115	- 0.059	0.053	1.02	0.4	0.43		
BK1116	-0.222	0.051	1.03	0.6	0.45		
BK1131	0.104	0.054	1.03	0.5	0.38		
BK1132	-0.335	0.052	0.93	- 1.3	0.46		
BK1133	-0.427	0.062	0.94	- 0.5	0.39		
BK1141	0.094	0.053	1.13	2.0	0.29		
BK1142	-0.072	0.051	1.07	1.2	0.34		
BK1143	- 0.954	0.055	1.01	0.2	0.44		
BK1212	0.084	0.062	0.90	- <b>2.4</b>	0.43		
BK1213	- 0.546	0.056	1.01	0.2	0.34		
BK1221	0.554	0.056	1.07	0.9	0.26		
BK1222	0.601	0.056	1.03	0.4	0.32		
BK1223	- 0.246	0.051	1.02	0.3	0.48		
BK1225	- 0.159	0.052	1.04	0.7	0.38		
BK1231	0.266	0.053	1.00	0	0.39		
BK1232	0.173	0.053	0.97	- 0.5	0.48		
BK1233	0.439	0.054	0.98	- 0.2	0.37		
BK1341	0.300	0.054	0.97	- 0.4	0.42		
BK1342	0.245	0.053	0.94	- 0.9	0.49		
BK1343	- 0.089	0.052	1.02	0.3	0.41		
BK1351	0.750	0.057	0.93	- 0.7	0.47		
BK1352	0.044	0.052	0.97	- 0.5	0.46		
BK1353	0.097	0.053	0.96	- 0.5	0.49		
BK14241	0.255	0.052	0.84	- <b>2.2</b>	0.55		
BK14243	- 0.358	0.050	0.97	- 0.5	0.5		
BK14244	- 0.127 <sup>a</sup>	0.290	0.95	- 0.9	0.54		

Table 11 Item information for one-dimensional partial-credit-model (after item exclusion)

Separation Reliability = 0.983; Chi-square test of parameter equality = 1630.66, df = 29, Sig. = 0.000

Problematic values are printed in bold

<sup>a</sup> The item parameters were fixed to zero as part of the item analysis. One parameter is fixed by ConQuest by default for model identification purposes

Tab	e 12	Glo	bal	DIF	anal	ysis	to	identif	y su	bgrou	ıp i	invariar	ice	amo	ong	trai	ning	courses
-----	------	-----	-----	-----	------	------	----	---------	------	-------	------	----------	-----	-----	-----	------	------	---------

	Estimate	Standard error
Nursing specialist	- 0.165	0.112
Geriatric nursing	- 0.07	0.12
Clinical nursing	0.235ª	
Chi-square of parameter equality	2.51	
df	2	
Sig Level	0.286	

<sup>a</sup> The item parameters were fixed to zero as part of the item analysis. One parameter is fixed by ConQuest by default for model identification purposes

ltem	Estimate (Nursing specialist)	Standard error	Estimate (Geriatric nursing)	Standard error	Estimate (Clinical nursing)		
BK1111	- 0.094	0.137	0.030	0.109	0.064 <sup>a</sup>		
BK1112	- 0.009	0.194	-0.107	0.146	0.117 <sup>a</sup>		
BK1113	0.008	0.133	0.035	0.109	$-0.043^{a}$		
BK1114	0.076	0.139	- 0.068	0.105	$-0.008^{a}$		
BK1115	-0.049	0.133	0.078	0.108	$-0.030^{a}$		
BK1116	-0.067	0.124	0.032	0.101	0.035 <sup>a</sup>		
BK1131	0.042	0.157	- 0.02	0.118	-0.021 <sup>a</sup>		
BK1132	-0.124	0.129	- 0.048	0.101	0.173 <sup>a</sup>		
BK1133	-0.015	0.220	0.389	0.186	- 0.373 <sup>a</sup>		
BK1141	0.251	0.162	0.004	0.118	- 0.255 <sup>a</sup>		
BK1142	0.036	0.132	- 0.022	0.102	$-0.014^{a}$		
BK1143	-0.204	0.158	0.102	0.119	0.101 <sup>a</sup>		
BK1212	- 0.559	0.195	0.435	0.166	<b>0.124</b> <sup>a</sup>		
BK1213	- 0.058	0.150	0.094	0.124	$-0.037^{a}$		
BK1221	0.004	0.179	- 0.092	0.137	0.088 <sup>a</sup>		
BK1222	-0.150	0.152	- 0.083	0.124	0.233 <sup>a</sup>		
BK1223	0.060	0.128	0.080	0.100	$-0.139^{a}$		
BK1225	-0.019	0.136	- 0.068	0.105	0.088 <sup>a</sup>		
BK1231	-0.187	0.133	- 0.053	0.107	0.239 <sup>a</sup>		
BK1232	-0.009	0.139	- 0.125	0.108	0.135 <sup>a</sup>		
BK1233	0.142	0.162	- 0.044	0.121	$-0.098^{a}$		
BK1341	0.172	0.175	- 0.087	0.126	$-0.086^{a}$		
BK1342	0.134	0.153	- 0.135	0.112	0.001 <sup>a</sup>		
BK1343	-0.040	0.129	-0.129	0.103	0.169 <sup>a</sup>		
BK1351	-0.066	0.179	0.086	0.149	$-0.020^{a}$		
BK1352	0.138	0.141	-0.112	0.106	$-0.026^{a}$		
BK1353	0.138	0.149	0.074	0.115	- 0.212 <sup>a</sup>		
BK14241	-0.100	0.131	0.061	0.106	0.039 <sup>a</sup>		
BK14243	0.262	0.127	0.018	0.099	$-0.279^{a}$		
BK14244	0.287 <sup>a</sup>		$-0.325^{a}$		0.038 <sup>a</sup>		

Separation Reliability = 0.118; Chi-square test of parameter equality = 51.29, df = 58, Sig. = 0.721

Significant values are printed in bold

<sup>a</sup> The item parameters were fixed to zero as part of the item analysis. One parameter is fixed by ConQuest by default for model identification purposes

#### Abbreviations

CFA Confirmatory Factor Analysis

CopeCo-N Stress Coping Competence instrument for Nursing

- DIF Differential Item Functioning
- ECD Evidence Centered Design
- SJT Situational Judgement Test
- wMNSQ weighted MeaN SQuare

## Acknowledgements

Not applicable.

#### Author contributions

JW, UW, SS and EW designed the research study. JW, WV, UW and LW designed the CopCo-N. SS and PK provided the data analyses. WV, PK, LW and AS conducted the data collection. JW, WV, SS and PK wrote this manuscript. All authors read and approved the final manuscript.

#### Funding

This research has been funded by the German Federal Ministry of Education and Research (21AP006).

#### Availability of data and materials

Data will be accessible through a research data center upon completion of the research project.

#### Declaration

#### **Competing interests**

The authors declare that they have no competing interests.

Received: 31 October 2022 Accepted: 7 June 2023 Published online: 22 June 2023

#### References

- Adams T, Anderson L, Turner M, Armstrong J (2011) Coping through a disaster: lessons from Hurricane Katrina. J Homel Secur Emerg Manage 8:1
- Böhle F, Weihrich M (2020) Das Konzept der Interaktionsarbeit [The concept of interaction work]. Zeitschrift Für Arbeitswissenschaft 74(1):9–22
- Böhn S, Deutscher VK (2021) Development and validation of a learning quality inventory for in-company training in VET (VET-LQI). Vocat Learn 14(1):23–53
- Bond TG, Yan Z, Heene M (2021) Applying the Rasch model. Fundamental Measurement in the Human Sciences. Routledge, New York
- Bundesministerium für Gesundheit (Federal Ministry of Health) (2022). Pflegeberufegesetz [Nursing Professions Act]. https://www.bundesgesundheitsministerium.de/pflegeberufegesetz.html
- Chan D, Schmitt N (2005) Situational-judgment-test. In: Evers A, Anderson N, Smit-Voskuijl O (eds) Handbook of personnel selection. Blackwell, London, pp 219–246
- de Meyrick J (2003) The Delphi method and health research. Health Educ 103(1):7–16
- Deutscher V, Winther E (2022) Assessment of vocational competences-definitions, issues and quality criteria. Research approaches on workplace learning: insights from a growing field. Springer, Cham, pp 305–320
- Flaxman PE, Bond FW (2010) Worksite stress management training: moderated effects and clinical significance. J Occup Health Psychol 15(4):347–358
- Frey A, Ruppert JJ (2013) Structuring and detecting competence. In: Beck K, Zlatkin-Troitschanskaia O (eds) From diagnostics to learning success. Proceedings in Vocational Education and Training. Sense Publishers, Rotterdam, pp 183–198
- Glaser J (2006) Arbeitsteilung, Pflegeorganisation und ganzheitliche Pflege—arbeitsorganisatorische Rahmenbedingungen für Interaktionsarbeit in der Pflege [Division of labor, nursing organization and holistic nursing work-organizational framework conditions for interaction work in nursing care]. In: Böhle F, Glaser J (ed) Arbeit in der Interaktion— Interaktion als Arbeit [Work in interaction-interaction as work]. VS Verlag für Sozialwissenschaften, pp 43–57.

Gnambs T, Nusser L (2019) The longitudinal measurement of reasoning abilities in students with special educational needs. Front Psycholog. https://doi.org/10.3389/fpsyg.2019.00232

- Hacker W, Reinhold S (1999) Beanspruchungsscreening bei Humandienstleistungen (BHD-System) [Stress Screening in Human Services]. Swets Test Services.
- Hacker W (2020) Psychische Regulation von Arbeitstätigkeiten 4.0 [Mental regulation of work activities 4.0], vol 51. vdf Hochschulverlag AG.
- Heier L, Gambashidze N, Hammerschmidt J, Riouchi D, Geiser F, Ernstmann N (2022) Development and testing of the situational judgement test to measure safety performance of healthcare professionals: an explorative cross-sectional study. Nurs Open 9(1):684–691
- Heininger SK, Baumgartner M, Zehner F, Burgkart R, Söllner N, Berberat PO, Gartmeier M (2021) Measuring hygiene competence: the picture-based situational judgement test HygiKo. BMC Med Educ 21(1):1–8
- Heinrichs M, Stächele T, Dormes G (2015) Stress und Stressbewältigung [Stress and Stress Management]. In: Hahlweg K, Hautzinger M, Margraf J, Rief W Fortschritte der Psychotherapie [Progess of Psychotherapy], vol 58. Göttingen.
- Hundenborn G (2007) Fallorientierte Didaktik in der Pfege Grundlagen und Beispiele für Ausbildung und Prüfung [Caseoriented didactics in nursing]. Urban & Fischer, Munich
- Isfort M, Rottländer R, Weidner F, Gehlen D, Hylla J, Tucman D (2018) Pflege-Thermometer 2018: Eine bundesweite Befragung von Leitungskräften zur Situation der Pflege und Patientenversorgung in der stationären Langzeitpflege in Deutschland [Nursing Thermometer 2018: A nationwide survey of managers on the situation of nursing and patient care in inpatient long-term care in Germany]. bfw. https://projekt-gravita.de/pflege-thermometer-2018/
- Jang KL, Thordarson DS, Stein MB, Cohan SL, Taylor S (2007) Coping styles and personality: A biometric analysis. Anxiety Stress Coping 20(1):17–24
- Joiko K, Schmauder M, Wolff G (2010) Psychische Belastung und Beanspruchung im Berufsleben: Erkennen Gestalten [Mental stress and strain in working life: Understanding – Shaping]. Ed: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin [Federal Institute for Occupational Safety and Health].
- Kaluza G (2011) Stressbewältigung: Trainingsmaterial zur psychologischen Gesundheitsförderung [Stress management: Training material for psychological health promotion], 2nd edn. Springer-Verlag, Berlin
- Kaspar R, Döring O, Wittmann E, Hartig J, Weyland U, Nauerth A, Möllers M, Rechenbach S, Simon J, Worofka I (2016) Competencies in geriatric nursing: empirical evidence from a computer-based large-scale assessment calibration study. Vocat Learn 9:185–206
- Kato T (2012) Development of the coping flexibility scale: evidence for the coping flexibility hypothesis. J Couns Psychol 59(2):262–273

Kato T (2015) Frequently used coping scales: a meta-analysis. Stress Health 31(4):315-323

Klotz V (2015) Diagnostik beruflicher Kompetenzentwicklung. Eine wirtschaftsdidaktische Modellierung für die kaufmännische Domäne [Diagnostics of vocational competence development. A business didactic modeling for the commercial domain]. Springer Fachmedien.

Krause A, Dorsemagen C (2007) Ergebnisse der Lehrerbelastungsforschung: Orientierung im Forschungsdschungel [Results of teacher stress research: orientation in the research jungle] https://www.researchgate.net/profile/Andre as\_Krause4/publication/226933557\_Ergebnisse\_der\_Lehrerbelastungsforschung\_Orientierung\_im\_Forschungs dschungel/links/54c9ebca0cf2f0b56c24f7c4/Ergebnisse-der-Lehrerbelastungsforschung-Orientierung-im-Forsc hungsdschungel.pdf

Lavoie P, Michaud C, Belisle M, Boyer L, Gosselin E, Grondin M, Larue C, Lavoie S, Pepin J (2018) Learning theories and tools for the assessment of core nursing competencies in simulation: a theoretical review. J Adv Nurs 74(2):239–250 Lazarus R, Folkman S (1984) Stress, appraisal, and coping. Springer, New York

- Lyon L (2011) Stress, coping, and health. A conceptual overview. In: Rice VH (ed) Handbook of stress, coping, and health: Implications for nursing research, theory, and practice. Sage Publications, Thousand Oaks, pp 2–21
- McDaniel MA, Nguyen NT (2001) Situational judgment tests: a review of practice and constructs assessed. Int J Sel Assess 9(1/2):103–113

Metz AM, Rothe HJ (2017) Screening psychischer Arbeitsbelastung: Ein Verfahren zur Gefährdungsbeurteilung [Screening of mental workload: a risk assessment procedure]. Springer, Berlin

Mislevy RJ (2013) Evidence-centered design for simulation-based assessment. Mil Med 178(10):107–114

Mulder M (2014) Conceptions of professional competence. In: Billett S, Harteis C, Gruber H, Eds. International handbook of research in professional and practice-based learning; pp 107–137

Mulder M, Weigel T, Collins K (2006) The concept of competence in the development of vocational education and training in selected EU member states. A critical analysis. J Vocat Educ Training 59(1):65–85

- Patterson F, Zibarras L, Ashworth V (2016) Situational judgement tests in medical education and training: research, theory and practice: AMEE Guide No. 100. Med Teacher 38(1):3–17
- Pohl S, Carstensen C (2012) NEPS technical report—Scaling the data of the competence tests (NEPS Working Paper No. 14). Otto-Friedrich-Universität, Nationales Bildungspanel, Bamberg. https://www.neps-data.de/Portals/0/Working% 20Papers/WP\_XIV.pdf. Accessed 04 Nov 2021.
- Reiser S, Schacht L, Thomm E, Figalist C, Janssen L, Schick K, Dörfle E, Berberat P, Gartmeier M, Bauer J (2022) A videobased situational judgement test of medical students' communication competence in patient encounters: development and first evaluation. Patient Educ Couns 105(5):1283–1289

Sarafis P, Rousaki E, Tsounis A, Malliarou M, Lahna L, Bamidis P, Niaklas D, Papastavrou E (2016) The impact of occupational stress on nurses caring behaviors and their health related quality of life. BMC Nurs 15:56

Schulze S, Holmberg C (2021) Bedeutung und Belastung von Pflegekräften während der Corona-Krise [Importance of and burden on nursing staff during the corona crisis]. Public Health Forum 29:1. https://doi.org/10.1515/ publef-2020-0114

Schuster N, Haun S, Hiller W (2011) Psychische Belastungen im Arbeitsalltag: Trainingsmanual zur Stärkung persönlicher Ressourcen [Mental stress in everyday work: training manual for strengthening personal resources]. Beltz Verlag.

Shavelson RJ (2012) Assessing business-planning competence using the Collegiate Learning Assessment as a prototype. Empir Res Vocat Educ Training 4(1):77–90

Smith AB, Rush R, Fallowfeld LJ, Velikova G, Sharpe M (2008) Rasch ft statistics and sample size considerations for polytomous data. BMC Med Res Methodol 8:33. https://doi.org/10.1186/1471-2288-8-33

Taylor BJ (2006) Factorial surveys: Using vignettes to study professional judgement. Br J Soc Work 36(7):1187–1207 Tuzinski K (2013) Simulations for personnel selection: an introduction. In: Fetzer M, Tuzinski K (eds) Simulations for personnel selection. Springer, Berlin, pp 1–13

Westermann C, Kozak A, Harling M, Nienhaus A (2014) Burnout intervention studies for inpatient elderly care nursing staff: systematic literature review. Int J Nurs Stud 51:63–71

Wittmann E, Weyland U, Seeber S, Warwas J, Strikovic A, Krebs P, Pohley M, Wilczek L (2022) Test sensitivity in assessing competencies in nursing education. Empir Res Vocat Educ Training 14:3. https://doi.org/10.1186/ s40461-022-00131-9

#### **Publisher's Note**

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