


RESEARCH ARTICLE

Validation of the instrument “Complexity of Nursing Care”—A mixed-methods study

Evelyn Huber¹  | Michael Kleinknecht-Dolf²  | Christiane Kugler³ | Marianne Müller⁴ | Rebecca Spirig^{1,2,5}

¹Department of Nursing Science, Faculty for Health, University Witten/Herdecke, Witten, Germany

²Department of Nursing and Allied Health Care Professionals, University Hospital Zurich, Zurich, Switzerland

³Faculty of Medicine, Institute of Nursing Science, University of Freiburg, Freiburg, Germany

⁴Institute of Data Analysis and Process Design, School of Engineering, Zurich University of Applied Sciences, Winterthur, Switzerland

⁵Department Public Health, Institute of Nursing Science, University of Basel, Basel, Switzerland

Correspondence

Michael Kleinknecht-Dolf, Universitätsspital Zürich, Rämistrasse 100/PHY 6, Zürich CH-8091, Switzerland.
Email: michael.kleinknecht@usz.ch

Funding information

The Nursing Science Foundation Switzerland, the Swiss Nurses Association, and another Foundation, that wishes to remain anonymous, supported this research programme.

Abstract

Aims: This study aimed to psychometrically test the instrument “Complexity of Nursing Care” and to broaden the understanding of the instrument's psychometrics and applicability.

Design: Embedded mixed-methods design.

Methods: We performed a cross-sectional study assessing all stationary patients of five Swiss hospitals daily for up to 5 days with the instrument “Complexity of Nursing Care” over a 1-month period in 2015. The scale's psychometrics were analysed using partial least square structural equation modelling. In the qualitative study section, we completed 12 case studies and analysed them case-wise and across cases. Quantitative and qualitative results were synthesized in tables.

Results: Structural equation modelling confirmed a reflective-formative second-order model of the instrument with good psychometric properties leading to a formula for the calculation of a complexity score. Qualitative results evolved descriptions of low and high extent of complexity. Narrative considerations of two raters deepened the understanding of the inter-rater reliability.

KEYWORDS

collective case studies, complexity of nursing care, embedded mixed-methods design, instrument validation, partial least square structural equation modelling, psychometrics, reflective-formative second-order model

1 | INTRODUCTION

In recent years, the number of patients in need of more demanding nursing care in acute care hospitals has increased (Dharmarajan et al., 2016; Smith et al., 2018). Additionally, a decrease in length of stay has been observed in many countries, leading to more complex nursing care situations on the hospital unit level (Bartholomeyczik, 2011; Borghans, Heijink, Kool, Lagoe, & Westert, 2008; Hamada, Sekimoto,

& Imanaka, 2012). This may jeopardize patient safety without suitable staffing (Aiken et al., 2017; Driscoll et al., 2018). For example, a recent study showed higher probability of death and length of stay and lower probability of being discharged home in internal medical patients screened as experiencing highly complex care episodes during their hospital stay (Bandini et al., 2018). Moreover, Registered Nurses (RN) at acute care hospitals felt overwhelmed in highly complex nursing situations if contextual and individual resources were

Müller and Spirig shared last-authorship.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. *Nursing Open* published by John Wiley & Sons Ltd.

not supportive (Kentischer, Kleinknecht-Dolf, Spirig, Frei, & Huber, 2018). This shows the need for an instrument for monitoring complexity of nursing care on the patient level in acute care hospitals to allocate adequate nursing staff resources according to grades of patient complexity, thus safeguarding high patient safety levels and keeping RNs motivated.

1.1 | Background

Complexity has been defined as different dynamically interacting components causing uncertainty, unpredictability and the need to find unique solutions (Kannampallil, Schauer, Cohen, & Patel, 2011; Lindberg & Lindberg, 2008). RNs require professional knowledge, skills, experience, attention, clinical reasoning abilities and supportive contextual conditions to be able to handle complex nursing care situations successfully (Hoeve, Brouwer, Roodbol, & Kunnen, 2018; Kentischer et al., 2018; Kuiper, O'Donnel, Pesut, & Turrise, 2017).

When we started developing our instrument in 2009, various attempts at defining and measuring complexity of nursing care had been made. Definitions found in the literature covered nurses' knowledge and skill levels about procedures and nursing care processes (Prescott & Phillips, 1988), patients' needs about education, decision-making and procedures (Fulton & Wilden, 1998), patients' nursing and medical conditions and the characteristics of nurses and care environments (Petryshen, Pallas, & Shamian, 1995) and nurses' education levels to generate changes and set priorities in patients' care (Blastorah et al., 2001). Of existing instruments, some measured complexity using instruments that also addressed other concepts (Berry, 1977; Fulton & Wilden, 1998; Prescott & Phillips, 1988). One instrument focussed on patients' needs rather than on demands on nursing care (Blastorah et al., 2001). Another followed an interprofessional approach (de Jonge et al., 2001).

Various nursing researchers developed their instruments relying on the "framework for the comparative analysis of the complexity of organizations" described by Perrow (1967). They assessed complexity for ambulatory care (Verran & Reid, 1987), nursing home care (Velasquez, 2007) or on the hospital unit level rather than on the patient level (Alexander & Kroposki, 2001; Leatt & Schneck, 1981; Overton, Schneck, & Hazlett, 1977). According to this framework, organizations differ about work complexity by the stability and understanding of the raw materials used and by the predictabilities or uncertainties concerning the technology workers apply to change the raw materials according to the organization's aims. In nursing, raw materials refer to patients and the stability of their health situations. Technology means the interventions RNs practice relating to the nursing care process. Additionally, work complexity diverges about variability in raw materials and technology, that is according to variability among patients and their relatives, health problems and nursing interventions. This framework defines complexity as a relational concept incorporating a patient's situation and the resultant demands on nursing care (Alexander & Kroposki, 2001; Kleinknecht-Dolf, Grand, et al., 2015; Overton et al., 1977; Perrow, 1967).

However, these instruments were not congruent with the purpose of our monitoring. Accordingly, we developed a new German

instrument following a framework for instrument development with the intersecting elements of an instrument's function, its content, its structure and its measurement model (Brühl, 2012). About its function, our instrument "Complexity of Nursing Care" is to be used for monitoring purposes and to differentiate between patients, wards and medical disciplines by means of quantifying complexity in one reliable value (Kleinknecht-Dolf, Grand, et al., 2015).

About the content, we relied on Perrow's (1967) framework, given its previous usefulness in nursing and composed this working definition: "Complexity of nursing care describes the demands on professional nursing in the acute care hospital setting. The extent of the complexity is determined by the degree of instability, variability and uncertainty that exists. These factors are influenced in part by the impact of the disease and its associated therapies on the patient and on the patient's everyday life. However, they are also influenced by decisions about care and interventions taken to deliver the best outcome for the patient" (Kleinknecht-Dolf, Grand, et al., 2015, pp. 592–593).

An instrument's structural model describes the relationships between latent variables, hence between constructs and sub-constructs that are not directly measurable such as complexity and its dimensions. The measurement model clarifies how the latent variable(s) are operationalized. Relationships are called reflective when measurement variables are effects of a latent variable and formative when they are their causes (Brühl, 2012; Hair, Hult, Ringle, & Sarstedt, 2017). The structural and measurement model of the first version of our instrument included 15 formative items (measurement variables) on three subscales (latent variables). Subscale One explored the complexity of the patient situation, Subscale Two the demands on nursing care and Subscale Three the required competencies of RNs. Clinical nurse specialists (CNS) discussed the instrument for face validity and congruence with Perrow's (1967) framework before it was investigated in a pilot study (Kleinknecht-Dolf, Grand, et al., 2015), followed by a sequential explanatory mixed-methods study between 2010–2014 (Kleinknecht-Dolf, Spichiger, et al., 2015). Cluster and regression analyses and focus group interviews favoured an instrument with seven items, collapsing some of the items, deleting items not statistically relevant and deleting Subscale Three (Guggenbühl, Hausmann, & Müller, 2012).

This advanced instrument is reevaluated with this study. Our aims were to psychometrically test the instrument "Complexity of Nursing Care", to represent complexity in one score and to broaden the understanding of the instrument's psychometrics and applicability.

2 | THE STUDY

2.1 | Design

We applied an embedded mixed-methods design with a leading quantitative study section (Creswell, 2014). The research protocol for this study was published elsewhere (Huber, Kleinknecht-Dolf, Müller, Kugler, & Spirig, 2016).

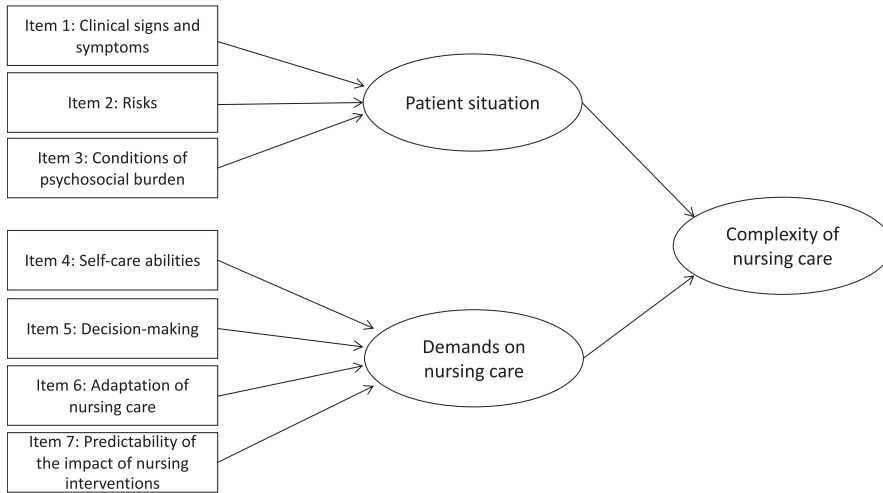


FIGURE 1 Preliminary formative-formative second-order model

2.2 | Methods

2.2.1 | Instrument

Based on Perrow's (1967) framework and the preceding results, we hypothesized a second-order measurement and structural model with the latent variables "patient situation" and "demands on nursing care" as first level and "complexity of nursing care" as second level and formative relationships between measurement and latent variables as shown in Figure 1 (Hair et al., 2017).

"Patient situation" was operationalized with: (a) clinical signs and symptoms; (b) risks; and (c) conditions of psychosocial burden. "Demands on nursing care" was measured with: (d) self-care abilities; (e) decision-making; (f) adaptation of nursing care; and (g) predictability of the impact of nursing interventions. The items were specified with a title, a description and examples of its meaning and contain a five-point Likert scale with an individual verbal description for every measurement point (Appendix S1). Furthermore, we provided examples for every measurement point in an additional form in four versions: for adult patients, paediatric patients, maternity wards and neonatal wards. Additionally, we added one global item for rating overall demands of the patient's care on RNs on a 10-point rating scale (1 = extremely low, 10 = extremely high).

2.2.2 | Quantitative data collection

The quantitative study section involved three Swiss university and two Swiss tertiary hospitals. In the prior pilot study, we tested the instrument "Complexity of Nursing Care" on six wards of two hospitals between April–May 2015. All new stationary patients were assessed daily by the RN responsible for the day shift for up to 5 days over a 2-week period. Some assessments were repeated by the head nurse, the CNS, or another RN on the ward for the estimation of the inter-rater reliability.

In the ensuing main cross-sectional study, all stationary patients were assessed over 4 weeks by the RN responsible for the day shift in all hospitals for up to 5 days starting from admission in November–December 2015 with the instrument "Complexity of

Nursing Care," added to the electronic documentation system in almost all wards.

2.2.3 | Qualitative data collection

The qualitative study section aimed to explore how complexity crystallized in various nursing care situations and how the instrument was applied in these situations. We performed 12 collective case studies between August–November 2015 (Creswell & Poth, 2018; Yin, 2014). Four nursing care situations were chosen in each of the following departments: (a) internal medicine; (b) cardiology, angiology and pneumatology; and (c) gynaecology. Pre-defined criteria per department included a patient with low complexity, one with high complexity, one with a hospital stay of 1–3 days and one with a hospital stay of at least 1 week.

For every situation, a researcher analysed the patient's record about information pertaining to our working definition on one specific day. On the same day, the RN primarily responsible for the patient and the ward's CNS assessed the situation independently with the instrument "Complexity of Nursing Care". Afterwards, they were interviewed in focussed single interviews (Flick, 2017). The interviews covered experiences while applying the instrument, considerations on how the situation was rated and perceptions of the patient's demands on nursing care. RNs employed for less than three months on the ward were not included.

2.2.4 | Syntheses

For the syntheses, the samples from the quantitative the qualitative study section were used.

2.3 | Data analysis

2.3.1 | Quantitative analyses

Inter-rater reliability

Double assessments from the pilot study were analysed on the item level using Spearman's rank correlation.

Model refinement

The preliminary hypothesized structural and measurement model was reviewed by calculating Gamma coefficients (γ) between the measurement variables and network analysis based on partial correlations to obtain information of the relationships among the measurement variables and thereby identify the actual model of the instrument to be evaluated (Costantini et al., 2015).

Evaluation of the structural and measurement model

We used partial least square structural equation modelling (PLS-SEM) for the evaluation of the refined structural and measurement model and the estimation of the instrument's psychometrics. PLS-SEM is an explorative modelling approach that aims to predict a target construct incorporating reflective and formative measures and thus to confirm construct validity of an instrument.

PLS-SEM is a non-parametric scaling method that does not require normal distribution of the data. Reflective models are evaluated by calculating indicator reliability and internal consistency of the measurement variables, convergent validity in terms of shared variance between the measurement variables of a reflective latent variable and discriminant validity defined as sufficient distinction between reflective latent variables. In formative models, collinearity and path coefficients are calculated to investigate whether the formative indicators do not measure similarly and that each indicator contributes sufficiently to the latent variable (Hair et al., 2017). The global item was added to estimate convergent validity of the formative model.

Statistical analyses were performed with IBM SPSS Statistics 23 (SPSS Inc, Released 2015) and the statistical computing language R 3.1.3 (R Foundation for Statistical Computing, 2016).

2.3.2 | Qualitative analyses

Within-case analysis

We performed within-case analysis inductively using summarizing and explicating content analysis before constructing a case narrative for each case (Mayring, 2015; Patton, 2015; Yin, 2014). MAXQDA12 was used in this step (MAXQDA, 1989–2016).

Cross-case analysis

Cross-case analysis was completed using a table applying structuring content analysis to explore similarities and differences among the cases' contents representing the dimensions instability, uncertainty and variability (Creswell & Poth, 2018; Mayring, 2015; Yin, 2014). Additionally, we compared the individual interviews about characteristics of different complexity levels. Statements concerning the instrument's applicability were analysed with summarizing content analysis (Mayring, 2015).

2.3.3 | Syntheses

We synthesized quantitative and qualitative results in three steps (Creswell, 2015; Fetters, Curry, & Creswell, 2013). First,

inter-rater reliability scores of the pilot study were contrasted in a table to those of the case studies and to narrative agreements and disagreements of the two interview participants of the same case. Second, in a discussion, the instrument's psychometrics were supplemented with the interview participants' experiences while applying the instrument. Third, the meaning of the complexity score in practice was explored by supplementing it with the cross-case results in a table.

2.4 | Ethics

A cantonal ethics committee assessed this research programme, which did not fall under Swiss human research legislation, as ethically unproblematic (Waiver No. 82/14 from 5 December 2014 and Waiver No. 49 from 22 May 2015). We followed Swiss legislation on data protection and guidelines for good clinical practice. All patient identification numbers and names of interview participants were replaced with codes. Participation in the interviews was voluntary. Interview participants signed the informed consent form after being given written and verbal study information by one of the researchers. All RNs involved in the quantitative data collection were given verbal and written study information and on-going support on the wards from a member of the research team. According to Swiss law, there was no requirement for asking consent from the patients or from RNs participating in the quantitative data collection. Data were stored safely.

2.5 | Rigour

We used the same theoretical understanding in every analytical step (Creswell, 2014). An expert statistician processed the statistical analyses. The first author reflected on analytical steps and results regularly with research team members and in peer groups (Rettke, Pretto, Spichiger, Frei, & Spirig, 2018). Further, we sent the case study results to the interview participants and three research team members for validation in Spring 2017 leading to adjustment of the results (Mayring, 2015).

3 | RESULTS

3.1 | Quantitative study section

In the pilot study, 748 assessments were performed on 246 patients. From these, 60 assessments of 47 patients were repeated by a second-rater. In the main study, 10,892 patients were assessed with 32,610 assessments on 189 wards to investigate the instrument in a broad variety of medical disciplines.

3.1.1 | Inter-rater reliability

The ratings ($N = 60$) of two pairs of raters showed moderate to strong positive correlations for all of the items with Spearman's Rho (r_s) between .703–.855 (Table 3).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Item 1	1.00	.76	.56	.68	.58	.69	.69
Item 2	.76	1.00	.64	.69	.64	.72	.76
Item 3	.56	.64	1.00	.53	.58	.59	.62
Item 4	.68	.69	.53	1.00	.79	.74	.67
Item 5	.58	.64	.58	.79	1.00	.71	.68
Item 6	.69	.72	.59	.74	.71	1.00	.78
Item 7	.69	.76	.62	.67	.68	.78	1.00

TABLE 1 Inter-item correlations based on gamma-coefficients, main study (N = 2,412)

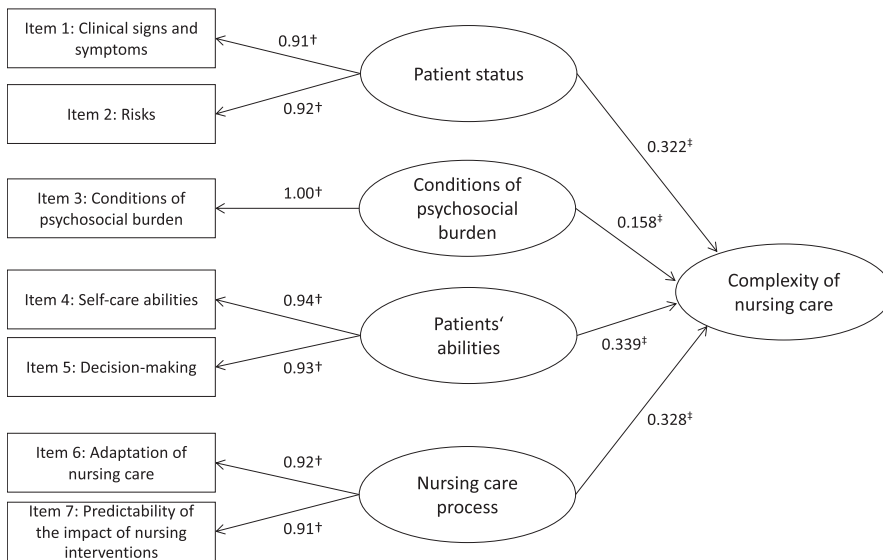


FIGURE 2 Reflective-formative second-order model
[†] Loadings of the reflective measurement model of the main study
[‡] Path coefficients of the formative structural model of the main study

3.1.2 | Model refinement

We included every patient's first assessment in the pilot study (N = 246) to refine the structural and measurement model. The first assessments of the patients of one university hospital were selected in the main study (N = 2,412) to review this refined model.

Items 1 and 2, items 4 and 5 and items 6 and 7 strongly correlated among each other in the pilot and the main study. Other correlations were slightly weaker (Table 1). Thus, these groups were assumed to build three formative indicators of the latent construct “complexity

of nursing care,” while item three proved to be an indicator of its own. Network analyses confirmed this model in the pilot and the main study. Accordingly, a reflective-formative second-order model was developed in the pilot study and was confirmed in the main study (Figure 2).

3.1.3 | Evaluation of the structural and measurement model

In the main study (N = 2,412), the reflective measurement model was supported through satisfactory estimates of:

TABLE 2 Discriminant validity, main study (N = 2,412)

Item	First-order latent variable	Patient status	Psycho-social burden	Patient's abilities	Nursing care process
1 Signs and symptoms	Patient status	0.91	0.52	0.62	0.63
2 Risks	Patient status	0.92	0.58	0.67	0.70
3 Psychosocial burden	Psychosocial burden	0.60	1.00	0.55	0.59
4 Self-care abilities	Patient's abilities	0.71	0.52	0.94	0.69
5 Decision-making	Patient's abilities	0.61	0.51	0.93	0.67
6 Adaptation nursing care	Nursing care process	0.68	0.54	0.70	0.92
7 Predictability of impact	Nursing care process	0.66	0.53	0.64	0.91

TABLE 3 Side-by-side matrix inter-rater reliability

Item	r_s pilot study	r_s case studies	Summary of the ratings of the case studies	Examples of narrative argumentations with numeric disagreements	Examples of narrative argumentations with numeric agreements	Interpretation
Clinical signs and symptoms	.713	.856	7 cases with identical ratings, 5 cases with differences between 0.5–1 point	<p><i>Consistent argumentation:</i> Case 3, RN (4)^a and CNS (5): multiple progressive signs and symptoms</p> <p><i>Non-consistent argumentation:</i> Case 2, RN (3): multiple, severe signs and symptoms, to some degree controllable, but always a topic of discussion. CNS (4): stable vital signs, severe, insufficiently controlled pain</p>	<p><i>Consistent argumentation:</i> Case 6, RN (2) and CNS (2): multiple, decreasing signs and symptoms.</p> <p><i>Non-consistent argumentation:</i> Case 9, RN (2): no discomfort at rest in spite of instable vital signs. CNS (2): typical clinical signs, familiar knowledge needed on basic education level</p>	Item 1 shows good inter-rater reliability despite certain subjectivity in the perception of the situation and some disagreement of how to rate the item
Risks	.758	.555	4 cases with identical ratings, 7 cases with a 1-point difference and 1 case with a 2-point difference	<p><i>Consistent argumentation:</i> Case 5, RN (3) and CNS (4): multiple moderate to high risks</p> <p><i>Non-consistent argumentation:</i> Case 7, RN (2): one risk to be controlled with an intervention being started. Possibility of hidden risks. CNS (4): multiple risks to monitor</p>	<p><i>Consistent argumentation:</i> Case 3, RN (5) and CNS (5): multiple, extremely severe risks with a risk of lethal deterioration.</p> <p><i>Non-consistent argumentation:</i> Case 1, RN (2): risks because of overweight. CNS (2): multiple risks because of actual primary disease and psychiatric co-morbidities</p>	Item 2 shows good inter-rater reliability in the pilot study. In the case studies, some cases indicate no consistent understanding of the item or of the situation and some disagreement of how to rate the item
Conditions of psychosocial burden	.772	.789	4 cases with identical ratings, 7 cases with differences between 0.5–1 point. 1 case with one missing	<p><i>Consistent argumentation:</i> Case 12, RN (4) and CNS (5): extremely burdened family with two family members on intensive care units</p> <p><i>Non-consistent argumentation:</i> Case 7, RN (2): patient shows no signs of insecurity. CNS (3): patient with personal resources and some signs of insecurity</p>	<p><i>Consistent argumentation:</i> Case 4, RN (2) and CNS (2): patient initially exhausted.</p> <p>Case 6, RN (2) and CNS (2): patient with personal and familial resources</p>	Item 3 shows good inter-rater reliability despite certain subjectivity in the perception of the situation and some disagreement of how to rate the item
Self-care abilities	.854	.781	6 cases with identical ratings, 4 cases with a 1-point difference, 1 case with a 2-point difference and 1 person who was undecided	<p><i>Consistent argumentation:</i> Case 2, RN (4) and CNS (5): extensive instructions and extensive support of ambulation needed, patient able to eat independently</p> <p><i>Non-consistent argumentation:</i> Case 9, RN (4): patient needs support and education. CNS (2): independent patient with unclear familiar support about his chronic conditions</p>	<p><i>Consistent argumentation:</i> Case 10, RN (5) and CNS (5): very passive patient.</p> <p><i>Non-consistent argumentation:</i> Case 1, RN (3): patient independent in personal hygiene, dependent in signing papers. CNS (3): patient partly dependent in personal hygiene with urinal incontinence and ambulation</p>	Item 4 shows good inter-rater reliability despite certain subjectivity in the perception of the situation and the argumentation and some disagreement on how to rate the item

(Continues)

TABLE 3 (Continued)

Item	r_s pilot study	r_s case studies	Summary of the ratings of the case studies	Examples of narrative argumentations with numeric disagreements	Examples of narrative argumentations with numeric agreements	Interpretation
Decision-making	.758	.866	6 cases with identical ratings, 5 cases with a 1-point difference, 1 person did not rate this item	<p><i>Non-consistent argumentation:</i> Case 11, RN (3): patient not capable of making decisions, relative needs support. CNS (2) patient not capable of making decisions, relative pretty much capable of making decisions</p>	<p><i>Consistent argumentation:</i> Case 10, RN (5) and CNS (5): patient not capable of making decisions, relative decides against professional recommendations. <i>Non-consistent argumentation:</i> Case 5, RN (3): patient knows her needs and concerns well, needs information for decisions. CNS (3): patient does not speak up for herself</p>	Item 5 shows good inter-rater reliability despite certain subjectivity in the perception of the situation and the argumentation and some uncertainty of how to rate the item
Adaptation of nursing care	.690	.683	6 cases with identical ratings, 4 cases with a 1-point difference and 2 cases with a 2 point difference	<p><i>Consistent argumentation:</i> Case 10, RN (2): situation more or less stable. CNS (3): situation assessable and controllable, no sudden deterioration expected. <i>Non-consistent argumentation:</i> Case 8, RN (1): the actual day only interventions realized that had been planned the day before. CNS (3): constant re-evaluations needed because of the symptoms with some routinized adaptations</p>	<p><i>Consistent argumentation:</i> Case 4, RN (1) and CNS (1): only routine interventions and monitoring. Case 12, RN (5) and CNS (5): constant adaptation because of high instability. <i>Non-consistent argumentation:</i> Case 11, RN (3): interventions have to be adapted to the patient's condition. CNS (3): risks have to be evaluated daily</p>	Item 6 shows acceptable inter-rater reliability, however, there is no consistent understanding of this item in some cases and some disagreement on how to rate the item. One person did not understand this item at all
Predictability of the impact of nursing interventions	.713	.585	7 cases with identical ratings, 4 cases with a 1-point difference and 1 case with a 3-point difference	<p><i>Non-consistent argumentation:</i> Case 7, RN (2): stabilized situation because of interventions taken. CNS (3): we know the direction, how the situation will develop, which side effects are to be expected, but we do not know exactly what will happen</p>	<p><i>Non-consistent argumentation:</i> Case 5, RN (3): this morning's interventions were effective. CNS (3): situation generally predictable, but still not too predictable because of the patient's multi-morbidity Case 8, RN (2): interventions about own practice are predictable; interventions outside own practice are not predictable. CNS (2): impacts of medications and conversations are mostly predictable</p>	Item 7 shows rather a good inter-rater reliability in the pilot study. In the case studies, some cases indicate no consistent understanding of this item and some disagreement on how to rate the item

Abbreviation: r_s , Spearman's rho.

^aThe number indicates the rating on the scale between 1–5 points, where every measurement point of the scale has an individual verbal description.

- indicator reliability: all first-order latent variables explained sufficient variance of the items (loadings reported in Figure 2). To represent the measurement model for complexity, the items were assigned to complexity in the form of the repeated indicators approach (loadings: item 1 = 0.79; item 2 = 0.84; item 3 = 0.73; item 4 = 0.85; item 5 = 0.81; item 6 = 0.84; item 7 = 0.81);
- internal consistency: Cronbach's alpha of items 1 and 2 about "patient status" was .80, of items 4 and 5 for "patient's abilities" .86 and of items 6 and 7 for "nursing care process" .81; Cronbach's alpha of all the seven items according to "complexity of nursing care" was .91;
- convergent validity: average variance extracted (AVE) confirmed convergence of the items about the suitable latent first-order variables with AVE-values of 0.83 for "patient status," 0.88 for "patient's abilities" and 0.84 for "nursing care process,"
- discriminant validity: four confined factors were approved (values reported in Table 2).

The latent variable "Conditions of psychosocial burden" contains only one item; thus, Cronbach's alpha and AVE equalled 1.00. Evaluations of the formative structural model confirmed the proposed model with estimates of:

- variance inflation factors (VIF) supporting the first-order latent variables to be independent formative indicators according to the second-order latent variable "complexity of nursing care" (VIF-value for "patient status": 2.64, for "conditions of psychosocial burden": 1.71, for "participation": 2.50 and for "nursing care process": 2.75);
- statistically significant path coefficients for all four first-order latent variables (weights reported in Figure 2) and for the global item (path coefficient: .7791).

Based on these results, a complexity score (the linear combination of the values of the seven items with multiplying factors derived from the measurement and structural model) can be calculated. This is in conjunction with a scaling factor to get complexity scores between 1.0–5.0, where 1.0 equals lowest complexity and 5.0 corresponds to highest complexity. Descriptive analyses, including all five hospitals, showed a right-skewed distribution of the complexity scores (first hospitalization day (1st hd), $N = 10,363$: mean (m) = 2.118, standard deviation (SD) = 0.814; 2nd hd, $N = 8,460$: $m = 2.145$, $SD = 0.807$; 3rd hd, $N = 6,379$: $m = 2.166$, $SD = 0.812$; 4th hd, $N = 4,435$: $m = 2.231$, $SD = 0.806$; 5th hd, $N = 2,545$: $m = 2.271$, $SD = 0.810$).

3.2 | Qualitative study section

The participants were between 24–51 years old (mean: 37 years) and had numerous years of professional experience (6 participants >20 years, 4 participants 11–15 years, 8 participants 6–10 years and 3 participants 1–5 years). The average interview length was 49 min (range: 26–72 min). Three CNSs participated in two case studies.

3.2.1 | Applicability of the instrument

Most interviewees had to become familiarized with the instrument before being able to apply it: "I had to read the questionnaire twice. I had to read it carefully. And while completing it, I still had some difficulties (...) even though there are very good examples, it is still somehow subjective," C10RN. Only one CNS judged the instrument as "easily understandable." Multiple patient problems with different severity levels, unstable fluctuating situations, psychosocial issues, patients who were unable to communicate, or autonomous patients in need of information caused uncertainty. Items 6 and 7 caused the most problems, either because participants did not understand their meaning or because of the need for constant adaptations in highly unstable situations. Overall, the interview participants were able to explain their ratings.

3.2.2 | Within- and cross-case results

Two examples of central characteristics of case narratives are shown in Table 4. Further case narratives have been published previously (Gurtner, Spirig, Staudacher, & Huber, 2018).

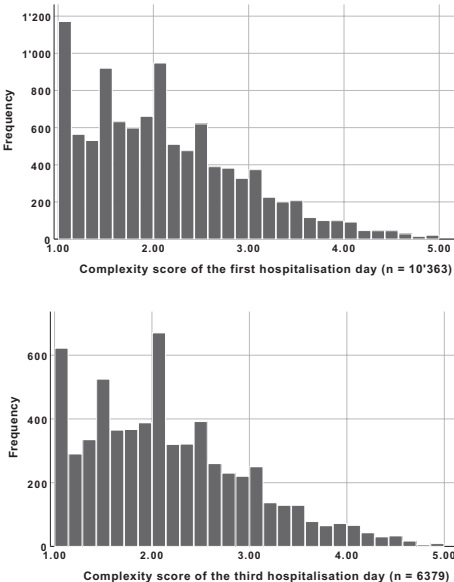
In cross-case analysis, patient-related complexity of nursing care in acute care hospitals was confirmed as dynamic phenomenon between low and high complexity in the dimension's instability, variability and uncertainty: "Slightly complex is close to complex. It can rapidly topple to the other side if the patient develops some complications," C9CNS. The extent of complexity appeared:

- in the type, number, assessability, controllability, progress and threat of nursing-relevant illness-related, therapy-related, psychosocial and ethical problems of patients and their relatives as well as the interactions between them: "the assessment is very high because of the sum of his risks," C12RN
- in the degree of concentrated attention, knowledge, experience and caring required from RNs to achieve an effective nursing care process for strengthening abilities and preventing deterioration, persistent harm and suffering of patients and for strengthening abilities and relieving distress of patients' relatives and thus delivering the best outcomes for patients and their relatives: "I sat next to him and observed him – I even almost studied him," C11RN.

Personal, cognitive, physical, functional, communicative and social resources of patients and their relatives were constitutive for the extent of complexity: "He has a long list of problems. But we have the impression that he is not suffering (...). He has a positive attitude; he is very thankful and patient," C6CNS.

The comparison of the interviews in ascending order concerning the theoretical dimensions instability, uncertainty and variability showed that low complexity was characterized as "nursing care situations requiring attention in routine nursing care interventions for patients with high resources in controlled, stable conditions." In contrast, high complexity was described as "nursing care situations requiring constant attention on severe, poorly assessable

TABLE 4 Distribution of the complexity score of the first and the third hospitalization day of all participating hospitals ($N = 10,363$ resp. 6,379) with a verbal description and an example of slightly and highly complex situations derived from the case studies

Slightly complex situations	Histograms (1.00 indicating lowest and 5.00 highest level of complexity)	Highly complex situations
<p>Nursing care situations requiring attention to routine nursing care interventions for patients with extensive resources in controlled stable conditions.</p> <p>Example Case 4: Independent patient with diverse internal and external resources and a previous experience with a similar hospitalization in an assessable and controllable condition in need of routine nursing interventions.</p> <p>C4CNS: "Here, we really practise routine interventions," complexity score: 1.425. C4RN: "She didn't need any support, we didn't need to educate her concerning pain, for example, what we could offer her, or how we could support her feeling well (...). She decided independently that it was good for her (...) and that she felt very comfortable the way it was," complexity score: 1.418.</p>	 <p>Frequency</p> <p>Complexity score of the first hospitalisation day (n = 10'363)</p> <p>Frequency</p> <p>Complexity score of the third hospitalisation day (n = 6379)</p>	<p>Nursing care situations requiring constant attention to severe, poorly assessable illness-related and psychosocial problems of patients in uncontrolled, unstable conditions with uncertain outcomes and limited or in some cases missing resources.</p> <p>Example Case 12: Patient with almost no abilities, with known, but very severe and highly instable health problems in need of constant monitoring and permanent adaptation of nursing interventions to ensure survival and with highly burdened relatives including a second close family member, who was hospitalized in a different ICU.</p> <p>C12CNS: "Here, one needs diverse concepts in one's mind to care for the patient and his family appropriately," complexity score: 4.851. C12RN: "This is a very sick patient. We focused on adjustments the whole time. This means, we do not know at all what happens if we adjust [the therapies], where direction this all moves," complexity score: 4.862.</p>

illness-related and psychosocial problems of patients in uncontrolled unstable conditions, with uncertain outcomes and limited or in some cases missing resources" (Table 4).

3.3 | Syntheses

3.3.1 | Inter-rater reliability and considerations of two raters

Items 1, 3, 4 and 5 showed comparable values of inter-rater reliabilities in the pilot study and the case studies (r_s : .718–.866). The disagreements between two interview participants were minor in these items.

The ratings of items 2 and 7 diverged in the pilot study (r_s : .757 respective .728) and the case studies (r_s : .555 respective .585). In item 6, the values of the pilot study and the case studies were rather low (r_s : .703 respective .683). The narrative argumentations revealed divergent understandings of how these items should be rated (i.e. case 10, item 7, RN: "today, nothing significant will happen," rated as "impacts and consequences of nursing interventions are predictable," CNS: "we do not know where this whole situation is supposed to end," rated as "short-term impacts and consequences are partly predictable, medium-term are mostly not predictable."

In the case studies, some pairs of raters reasoned the same way but chose another rating in most of the items. Some raters reasoned differently about the same item but still arrived at an identical numerical rating. Finally, there was also consensus according to numerical rating and narrative explanations (Table 3).

3.3.2 | Construct validity and applicability

The psychometric evaluations of the instrument supported its construct validity and reliability with all items yielding comparably good values. However, the interviews revealed that items 6 and 7 in particular seemed to be more difficult to rate than the others. Furthermore, the application of this instrument needed reflection. It can be hypothesized that these uncertainties diminish if raters become more familiar with the instrument, or that the amount of quantitative data equalizes some uncertainties.

3.3.3 | Distribution and description of the complexity score

PLS-SEM led to a formula for the calculation of a continuous complexity score. Merging the complexity scores with cross-case results indicated that only a lower ratio of nursing care situations demanded high levels of attention, knowledge, skills and caring expertise to handle conditions of patients with limited abilities and multiple, instable, poorly controllable health, psychosocial and ethical problems (Table 4).

4 | DISCUSSION

With this study, we developed a new instrument with satisfactory psychometric properties. It was evaluated using an adequate number of assessments in a broad variety of acute care hospital disciplines. This instrument makes it possible to capture the complexity

of a nursing care situation in one reliable value. With only seven items, not much effort is needed to use this instrument in practice or research.

In the leading quantitative study section, we elaborated and evaluated a reflective-formative second-order model. All indices supported data-model fit and thus reliability and construct validity of the instrument sufficiently (Hair et al., 2017). Considering the first version of this instrument contained only formative indicators, this is an important advancement. With the former version, it was not possible to calculate a reliable composite score and to use the data for analyses other than descriptive evaluations based on ordinal increments. The advanced instrument allows us to work with the data by means of a continuous composite score and thus combine them with other monitoring data for more complex evaluations.

Interestingly, our preliminary operationalization of the sub-construct "patient situation" with the items "clinical signs and symptoms," "risks" and "conditions of psychosocial burden" was differentiated in such a way that "clinical signs and symptoms" and "risks" refer to the present clinical state of a patient and "conditions of psychosocial burden" point to a more holistic knowledge of the patient and his or her family. Also, the operationalization of the "demands on nursing care" appeared to be composed of two sub-constructs. "Patients abilities" describes demands on nursing care if patients' abilities are limited, while "nursing care process" represents the technical aspect of nursing care. This advanced our conceptual understanding based on Perrow's (1967) framework.

Additionally, the qualitative study section added valuable information to the instrument and the theoretical framework. In particular, the sub-construct "patient's abilities" illustrates the relational understanding of the complexity of nursing care. "Self-care abilities" and "decision-making" of patients or patients' relatives relieve nurses about "doing for" and "enabling" patients with limited abilities, two characteristics of the concept of caring (Swanson & Wojnar, 2004). Abilities of self-management or decision-making are examples of internal resources, which help to effectively adopt external resources like healthcare services or social support (Steverink, Lindenberg, & Slaets, 2005). Interview participants incorporated additional resources of patients and their relatives. For instance, familial social support reduced or, if lacking, increased the participant's ratings of psychosocial burden. Similarly, other authors described patients' internal and external resources as defining elements of patient complexity (Schaink et al., 2012; Shippee, Shah, May, Mair, & Montori, 2012).

Synthesis of the quantitative and qualitative study strands supports overvalue of an embedded mixed-methods design in instrument development. In particular, the conjunction of the quantitative measurements of inter-rater reliability with an analysis of considerations of two different raters yielded helpful explanations for marginally acceptable statistical values. This led us to rephrase the wording of some of the descriptions and examples in the instrument. Furthermore, the instrument's robustness has to be questioned in items where pairs of raters used different ratings while indicating the same arguments. This initiated revisions

of the verbal descriptions of the measurement points of some of the items. Also, the interview participants' answers concerning the instrument's applicability impelled improvements in the wording. With these advancements, we hope to further increase the instrument's inter-rater reliability and help nurses to apply our new instrument with greater security.

Furthermore, our aim was representing and interpreting the complexity of nursing care with one reliable value, which was achieved in the quantitative study strand. By means of the qualitative study strand, we were able to characterize low and high grades of complexity. Our theoretical understanding of complexity is based on Perrow's (1967) framework for the complexity of work in organizations. Accordingly, the extent of complexity depends on the patient's situation as well as on professional demands to change the situation in accordance with the patients' aims. Hence, the extent of the complexity of a nursing care situation is to be viewed by taking professionally trained RNs into account. Synthesis of quantitative and qualitative results showed that a smaller portion of patients required high levels of concentrated attention, knowledge, experience and caring from RNs. Nursing managers should support RNs in complex nursing care situations, as these situations can be perceived either as a challenge or as an overwhelming burden (Kentischer et al., 2018). Hence, this instrument may be of use not only to gain monitoring data, but also to red-flag highly complex nursing care situations demanding further attention and support. Future monitoring will reveal whether the number of patients in highly complex nursing care situations will increase.

To our knowledge, the use of an embedded mixed-methods design has not yet been reported in instrument development in the literature. We have chosen this approach where the reevaluation of the instrument took place in an advanced stage of its development. Earlier publications about instrument development based on mixed-methods research followed sequential designs to strengthen the instrument's validity either by defining items with the use of a first qualitative study (Shahbazi Sighaldehy et al., 2019) or by explaining quantitative findings with a second qualitative study (Kleinknecht-Dolf, Spichiger, Müller, Bartholomeyczik, & Spirig, 2017). In contrast, our intention was to comprehend and interpret the instrument's inter-rater reliability, construct validity, applicability and the meaning of the derived complexity score. The overvalue of the results from the synthesis of the quantitative and qualitative study sections supports this method as useful for future use in a monitoring setting.

4.1 | Limitations

However, this study has some limitations. In the quantitative study section, criterion validity could not be investigated because of a lack of instruments in line with this study's aims. Moreover, insights in the course of complexity over time and an adequate summarized representation of the complexity over the course of hospitalization were unexplored. About the qualitative study section, more cases can be assumed to generate more insights into the perception of complexity of nursing care in

practice. Also, it can be assumed that difficulties in the instrument's application limit its psychometric properties. Also, this instrument requires some time for reflection, which RNs may lack in day-to-day practice.

5 | CONCLUSIONS

In summary, we present a short instrument with evidence of acceptable reliability and validity for the monitoring of patient-related complexity of nursing care in acute care hospitals. However, we found four sub-constructs instead of the originally intended two in the evaluation of the proposed structural and measurement model. This expanded our understanding of patient-related complexity of nursing care based on Perrow's (1967) framework. Our data highlight that a newly developed instrument should operationalize the theoretical framework in the way that empirical data confirm and advance the theoretical understanding to be studied, which again helps to advance the instrument (Brühl, Planer, & Grebe, 2012).

However, the need to gain familiarity with this instrument shows the necessity of careful instructions and support. Considering the findings of this study may help to purposefully prepare RNs' applying this new instrument and thus further improve its understanding, applicability and psychometric properties. Additional research is needed to merge the values of the patient-related complexity of nursing care with other monitoring data such as workload, staffing levels, skill-grade mix, nursing outcomes and nursing-sensitive patient outcomes. The results of this study provide a basis for doing so.

ACKNOWLEDGEMENTS

We thank the steering and project members, all the involved nurses and IT technicians from the participating hospitals for their collaboration in the preparation and data collection of this study, Roger Vogt for his personal data management support and Jane White for proofreading the manuscript.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

EH, MK-D, MM, CK and RS developed the design of the study. MM, EH and MK-D were involved in the collection and management of the data. MM analysed the statistical data, EH performed the qualitative data analysis and the synthesis of the quantitative and qualitative results. EH, MK-D, MM and RS interpreted the results. EH drafted the manuscript. All authors critically revised the manuscript for important intellectual content and approved the final version of the article. EH works in this research programme as a PhD student focussing on "complexity of nursing care" at the University of Witten/Herdecke, Germany. RS and CK are EH's PhD advisors. RS is the sponsor and

MK-D the scientific and project leader of this research programme. MM is the leading statistician.

ORCID

Evelyn Huber  <https://orcid.org/0000-0002-1572-551X>

Michael Kleinknecht-Dolf  <https://orcid.org/0000-0002-4652-7743>

REFERENCES

- Aiken, L. H., Sloane, D., Griffiths, P., Rafferty, A. M., Bruyneel, L., McHugh, M. ... Consortium, R. C (2017). Nursing skill mix in European hospitals: Cross-sectional study of the association with mortality, patient ratings, and quality of care. *BMJ Quality & Safety*, 26(7), 559–568. <https://doi.org/10.1136/bmjqs-2016-005567>
- Alexander, J. W., & Kroposki, M. (2001). Using a management perspective to define and measure changes in nursing technology. *Journal of Advanced Nursing*, 35(5), 776–783. <https://doi.org/10.1046/j.1365-2648.2001.01910.x>
- Bandini, F., Guidi, S., Blaszczyk, S., Fumarulo, A., Pierini, M., Pratesi, P., ... Landini, G. (2018). Complexity in internal medicine wards: A novel screening method and implications for management. *Journal of Evaluation in Clinical Practise*, 24, 285–292. <https://doi.org/10.1111/jep.12875>
- Bartholomeyczik, S. (2011). [Faster throughput of sicker patients in hospitals: Where remains nursing?] *Ethik in der Medizin*, 23, 315–325. <https://doi.org/10.1007/s00481-011-0160-x>
- Berry, D. M. (1977). An inpatient classification system for nursing service staffing decisions. *Communicating Nursing Research*, 8, 90–100.
- Blastorah, M., Alvarado, K., Duhn, L., Flint, F., McGrath, P., & VanDeVelde-Coke, S. (2001). Development and evaluation of an RN/RPN utilization toolkit. *Nursing Leadership*, 23 Special Issue, 33–55.
- Borghans, I., Heijink, R., Kool, T., Lagoe, R. J., & Westert, G. P. (2008). Benchmarking and reducing length of stay in Dutch hospitals. *BMC Health Services Research*, 8, 220. <https://doi.org/10.1186/1472-6963-8-220>
- Brühl, A. (2012). [Methodological orientation of nursing science in the development of standardised measurement instruments]. In A. Brühl (Ed.), *[Measuring nursing care needs? Challenges in developing nursing measurement instruments using the example of the new assessment instrument NBA]* (pp. 13–49). Vallendar, Germany: Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung der Philosophisch-Theologischen Hochschule Vallendar.
- Brühl, A., Planer, K., & Grebe, C. (2012). [How are better standardised measurement instruments developed in nursing? New methodological approaches to measure needs of nursing care]. In A. Brühl (Ed.), *[Measuring nursing care needs? Challenges in developing nursing measurement instruments using the example of the new assessment instrument NBA]* (pp. 151–170). Vallendar, Germany: Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung der Philosophisch-Theologischen Hochschule Vallendar.
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L., & Cramer, A. (2015). State of the art personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 54, 13–29. <https://doi.org/10.1016/j.jrp.2014.07.003>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage Publications.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Thousand Oaks, CA: SAGE Publications.

- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry research design. Choosing among five approaches* (4th ed.). Thousand Oaks, CA: Sage Publications.
- de Jonge, P., Huyse, F. J., Slaets, J. P. J., Herzog, T., Lobo, A., Lyons, J. S., ... Mellenbergh, G. J. (2001). Care complexity in the general hospital: Results from a European study. *Psychosomatics*, 42(3), 204–212. <https://doi.org/10.1176/appi.psy.42.3.204>
- Dharmarajan, K., Strait, K. M., Tinetti, M. E., Lagu, T., Lindenauer, P. K., Lynn, J., ... Krumholz, H. M. (2016). Treatment for multiple acute cardiopulmonary conditions in older adults hospitalized with pneumonia, chronic obstructive pulmonary disease, or heart failure. *Journal of the American Geriatrics Society*, 64(8), 1574–1582. <https://doi.org/10.1111/jgs.14303>
- Driscoll, A., Grant, M. J., Carroll, D., Dalton, S., Deaton, C., Jones, I., ... Astin, F. (2018). The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: A systematic review and meta-analysis. *European Journal of Cardiovascular Nursing*, 17(1), 6–22. <https://doi.org/10.1177/1474515117721561>
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs-principles and practices. *Health Services Research*, 48(6 Pt 2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
- Flick, U. (2017). *[Qualitative social research]* (8th ed.). Reinbeck bei Hamburg, Germany: Rowolt Taschenbuch Verlag.
- Fulton, T. R., & Wilden, B. M. (1998). Patient requirements for nursing care: The development of an instrument. *Canadian Journal of Nursing Administration*, 11(1), 31–51.
- Guggenbühl, J., Hausmann, A., & Müller, M. (2012). *[Measuring complexity in acute care hospitals. Project work degree program industrial engineering, business arithmetics]*. Winterthur, Switzerland: School of Engineering, Institute of Data Analysis and Process Design, Zurich University of Applied Sciences.
- Gurtner, C., Spirig, R., Staudacher, D., & Huber, E. (2018). [Patient-related complexity in nursing care - collective case studies in the acute care hospital]. *Pflege*, 31(5), 237–244. <https://doi.org/10.1024/1012-5302/a000626>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Los Angeles, CA: Sage Publications.
- Hamada, H., Sekimoto, M., & Imanaka, Y. (2012). Effects of the per diem prospective payment system with DRG-like grouping system (DPC/PDPS) on resource usage and healthcare quality in Japan. *Health Policy*, 107(2–3), 194–201. <https://doi.org/10.1016/j.healthpol.2012.01.002>
- Hoeve, Y. T., Brouwer, J., Roodbol, P. F., & Kunnen, S. (2018). The importance of contextual, relational and cognitive factors for novice nurses' emotional state and affective commitment to the profession. A multilevel study. *Journal of Advanced Nursing*, 74, 2082–2093. <https://doi.org/10.1111/jan.13709>
- Huber, E., Kleinknecht-Dolf, M., Müller, M., Kugler, C., & Spirig, R. (2016). Mixed-method research protocol: Defining and operationalizing patient-related complexity of nursing care in acute care hospitals. *Journal of Advanced Nursing*, 73(6), 1491–1501. <https://doi.org/10.1111/jan.13218>
- Kannampallil, T. G., Schauer, G. F., Cohen, T., & Patel, V. L. (2011). Considering complexity in healthcare systems. *Journal of Biomedical Informatics*, 44(6), 943–947. <https://doi.org/10.1016/j.jbi.2011.06.006>
- Kentischer, F., Kleinknecht-Dolf, M., Spirig, R., Frei, I. A., & Huber, E. (2018). Patient-related complexity of care: A challenge or overwhelming burden for nurses - A qualitative study. *Scandinavian Journal of Caring Sciences*, 32, 204–212. <https://doi.org/10.1111/scs.12449>
- Kleinknecht-Dolf, M., Grand, F., Spichiger, E., Müller, M., Martin, J. S., & Spirig, R. (2015). Complexity of nursing care in acute care hospital patients: Results of a pilot study with a newly developed questionnaire. *Scandinavian Journal of Caring Sciences*, 29(3), 591–602. <https://doi.org/10.1111/scs.12180>
- Kleinknecht-Dolf, M., Spichiger, E., Frei, I. A., Müller, M., Martin, J. S., & Spirig, R. (2015). Monitoring of nursing service context factors: First descriptive results of a Swiss study in the context of the introduction of DRGs. *Pflege*, 28(2), 93–107. <https://doi.org/10.1024/1012-5302/a000411>
- Kleinknecht-Dolf, M., Spichiger, E., Müller, M., Bartholomeyczik, S., & Spirig, R. (2017). Advancement of the German version of the moral distress scale for acute care nurses-A mixed methods study. *Nursing Open*, 4(4), 251–266. <https://doi.org/10.1002/nop.2.91>
- Kuiper, R. A., O'Donnel, S., Pesut, D., & Turrise, S. (2017). *The essentials of clinical reasoning for nurses. Using the outcome-present state-test model for reflective practice*. Indianapolis, IN: Sigma Theta Tau International.
- Leatt, P., & Schneck, R. (1981). Nursing subunit technology: A replication. *Administrative Science Quarterly*, 26(2), 225–236. <https://doi.org/10.2307/2392470>
- Lindberg, C., & Lindberg, C. (2008). Nurses take note: A primer on complexity science. In C. Lindberg, S. Nash, & C. Lindberg (Eds.), *On the edge: Nursing in the age of complexity* (pp. 23–47). Bordentown, NJ: PlexusPress.
- MAXQDA (1989–2016). MAXQDA, software for qualitative data analyses. Retrieved from www.maxqda.de
- Mayring, P. (2015). *[Qualitative content analyses: Basics and technics]* (12th ed.). Weinheim, Germany; Basel, Switzerland: Beltz Verlag.
- Overton, P., Schneck, R., & Hazlett, C. B. (1977). An empirical study of the technology of nursing subunits. *Administrative Science Quarterly*, 22(2), 203–219. <https://doi.org/10.2307/2391956>
- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Los Angeles, CA: Sage Publications Ltd.
- Perrow, C. (1967). A framework for the comparative analysis of organizations. *American Sociological Review*, 32(2), 194–208. <https://doi.org/10.2307/2091811>
- Petryshen, P., Pallas, L. L., & Shamian, J. (1995). Outcomes monitoring: Adjusting for risk factors, severity of illness, and complexity of care. *Journal of the American Medical Informatics Association*, 2(4), 243–249. <https://doi.org/10.1136/jamia.1995.96010393>
- Prescott, P. A., & Phillips, C. Y. (1988). Gauging nursing intensity to bring costs to light. *Nursing and Health Care*, 9(1), 17–22.
- R Foundation for Statistical Computing (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rettke, H., Pretto, M., Spichiger, E., Frei, I. A., & Spirig, R. (2018). Using reflexive thinking to establish rigour in qualitative research. *Nursing Research*, 67, 490–497. <https://doi.org/10.1097/NNR.0000000000000307>
- Schank, A. K., Kuluski, K., Lyons, R. F., Fortin, M., Jadad, A. R., Upshur, R., & Wodchis, W. P. (2012). A scoping review and thematic classification of patient complexity: Offering a unifying framework. *Journal of Comorbidity*, 2, 1–9. <https://doi.org/10.15256/joc.2012.2.15>
- Shahbazi Sighaldehy, S., Baheiraei, A., Ebadi, A., Khaki, I., Kelishadi, R., & Majdzadeh, R. (2019). Development and psychometric properties of the Hookah Smoking Initiation for Women Questionnaire (HIWQ). *Health Promotion International*, 34, 462–468. <https://doi.org/10.1093/heapro/dax085>
- Shippee, N. D., Shah, N. D., May, C. R., Mair, F. S., & Montori, V. M. (2012). Cumulative complexity: A functional, patient-centered model of patient complexity can improve research and practice. *Journal of Clinical Epidemiology*, 65(10), 1041–1051. <https://doi.org/10.1016/j.jclinepi.2012.05.005>
- Smith, S. M., Wallace, E., Salisbury, C., Sasseville, M., Bayliss, E., & Fortin, M. (2018). A core outcome set for multimorbidity research (COSmm). *Annals of Family Medicine*, 16(2), 132–138. <https://doi.org/10.1370/afm.2178>

- SPSS Inc. (Released 2015). *IBM SPSS Statistics for Windows, Version 23*. Armonk, NY: IBM Corp. Retrieved from www.spss.com
- Steverink, N., Lindenberg, S., & Slaets, J. P. J. (2005). How to understand and improve older people's self-management of wellbeing. *European Journal of Ageing*, 2(4), 235–244. <https://doi.org/10.1007/s10433-005-0012-y>
- Swanson, K. M., & Wojnar, D. M. (2004). Optimal healing environments in nursing. *Journal of Alternative and Complementary Medicine*, 10(Suppl 1), S43–S48. <https://doi.org/10.1089/1075553042245935>
- Velasquez, D. (2007). The development and testing of a questionnaire to measure complexity of nursing work performed in nursing homes: NCCQ-NH. *Geriatric Nursing*, 28(2), 90–98. <https://doi.org/10.1016/j.gerinurse.2007.01.009>
- Verran, J. A., & Reid, P. J. (1987). Replicated testing of the nursing technology model. *Nursing Research*, 36(3), 190–194. <https://doi.org/10.1097/00006199-198705000-00024>
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Los Angeles, CA: SAGE Publications.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Huber E, Kleinknecht-Dolf M, Kugler C, Müller M, Spirig R. Validation of the instrument "Complexity of Nursing Care"—A mixed-methods study. *Nursing Open*. 2020;7:212–224. <https://doi.org/10.1002/nop2.383>