

Assessment of empathy by simulated patients: Adaptation and validation of a new instrument

Abstract

Background: Medical doctors' empathy is vital in their interaction with patients, impacting on patient trust and health outcomes. The Consultation and Relational Empathy (CARE) scale helps assess healthcare professionals' empathy but there is a reasonable doubt about its validity when directly implemented in simulation contexts. This study adapts and validates a version of the CARE scale for simulated patients (Sp-SIMCARE), filling a gap in empathy assessment within medical training simulations.

Methods: The contextual adaptation of the CARE scale for simulated patients followed a four-phase process: 1) preparation of a preliminary adaptation proposal; 2) drafting the first version; 3) piloting the initial version with simulated patients; and 4) refining the final Sp-SIMCARE version. A panel of five experts collaborated with five simulated patients to ensure clarity, relevance, and language equivalence. The validation was conducted using typical primary care scenarios. Simulated patients assessed the performance of undergraduate medical students in four primary patient roles: acute, chronic, with high functional component, and with a hostile attitude. Psychometric parameters that were evaluated included convergent validity (assessed by simulated patients using a global score), acceptability and face validity, homogeneity, and internal reliability.

Results: The adaptation process resulted in a clear, relevant, and comprehensible scale, ensuring uniform understanding among users. Validation involved 95 students in 270 encounters with eight simulated patients. The final version showed convergent validity (Spearman's $\rho=0.730$, $p<0.001$), acceptability and face validity (1.96% "Does not apply"/missing values), homogeneity (item-total correlations 0.705-0.865) and reliability (Cronbach's $\alpha=0.960$).

Conclusions: The Sp-SIMCARE scale appears to be a valid and reliable tool for simulated patients to assess empathy in future doctors during their training from a multidimensional perspective.

Keywords: empathy, medical education, medical student, patient simulation, validation study, psychometrics

Pedro Brotons de los Reyes^{1,2,3}
Montserrat Virumbrales Cancio¹
Pere Castellvi¹
Xavier Martínez-Regada^{1,4}
Albert Balaguer^{1,2,5}

1 Universitat Internacional de Catalunya, School of Medicine and Health Sciences, Department of Medicine, Sant Cugat del Vallès, Spain

2 Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain

3 Instituto de Salud Carlos III, Consorcio de Investigación Biomédica en Red de Epidemiología y Salud (CIBERESP), Madrid, Spain

4 EAP Amadeu Torner, L'Hospitalet de Llobregat, Barcelona, Spain

5 Hospital Universitari General de Catalunya, Department of Pediatrics, Barcelona, Spain

1. Introduction

Medical doctors' ability to empathize with their patients is essential for establishing trust-based and effective patient-doctor relationships, leading to enhanced clinical outcomes [1]. Despite the emphasis placed on its importance, there is a lack of consensus on the definition of empathy in the specific clinical context. Commonly it is described as a cognitive ability to understand a patient's inner experiences and perspective, along with the capacity to communicate this understanding [2]. Some authors have distinguished four key components in empathy: emotive (the ability to share the patient's emotional state), moral (the motivation to act empathetically), cognitive (the capacity to understand the patient's perspective), and behavioural (the ability to communicate and act in a helpful or therapeutic manner with the patient) [3].

Despite differing models, over recent years most researchers have converged on the importance of clinicians recognizing and expressing awareness of patients' emotional states through behavioural responses [4], [5], [6]. Social neuroscience has recently progressed in identifying that empathy components that can be modified through changes in experiences [7]. Therefore, empathy is a competency that can be taught and learned, especially in the early stages of medical practice. Training in this competence could enhance health care professionals' and students' ability to provide effective medical care, while improving patients' experiences and engagement [8], [9], [10]. Several instruments are available to assess

medical doctors' empathy within general practice contexts, such as the Davis' Interpersonal Reactivity Index (IRI), the Jefferson Scale of Physician Empathy (JSPE), and the Consultation and Relational Empathy (CARE) scale [11], among others. Particularly, the CARE scale evaluates

1. the ability to understand the patient's situation, perspective and feelings (and their attached meanings);
2. to communicate that understanding and check its accuracy; and
3. to act on that understanding with the patient in a helpful (therapeutic) way [12].

Although the IRI, the JSPE, and other analogous instruments possess pertinent characteristics for the assessment of the emotive and cognitive dimensions of empathy, the CARE scale is distinctive in that it also evaluates the subsequent behavioural response, which is of paramount importance for both patient care and the training of medical students and professionals [13].

The CARE scale has undergone translation and validation in primary and specialized care consultations throughout Europe, North America, and Asia [14], [15], [16], [17], [18]. In addition, it has been directly implemented for assessment of empathy in simulation contexts [19], [20], [21], [22], [23]. However, although it seems reasonable that an instrument that has good metric properties in the real context will maintain similar and adequate results in a simulated context, this hypothesis should be tested to ensure its validity.

The objective of this study was to adapt the Spanish version of the CARE instrument (Sp-CARE) [24] for use among undergraduate medical students in simulated settings and to evaluate its metric properties.

2. Methods

2.1. Design and study population

This study took place at the Faculty of Medicine and Health Sciences of the Universitat Internacional de Catalunya (UIC), located in Barcelona, Spain, during the academic course 2022-2023. The participants included fourth-year undergraduate medical students and SPs who had wide experience in the evaluation of "core clinical competencies", such as interpersonal, communication and listening skills, time management, problem-solving, leadership, and empathy, among others, using standardized empathy measurement tools. The study protocol was approved by the ethics committee at the study centre (registration MED-2022-07).

2.2. The consultation and relational empathy measure

The CARE measure was developed and validated in 2004, showing strong correlation with other empathy-related scales and high internal reliability (Cronbach's alpha,

0.93) [12]. The scale consists of 10 items scored on a Likert scale from "1" (poor) to "5" (excellent), with higher scores signifying greater levels of empathy within a range from 10 to 50. Developers of the measure suggest that items 1 to 6 are primarily related to the affective aspect of empathy, whereas the remaining items are related to cognitive and behavioural aspects. An option of "does not apply" is also available for each item and can be selected if the item is deemed irrelevant to the consultation. Thus, minimal occurrences of "does not apply" or missing responses demonstrate the patients' perceived relevance of the item. To ensure practical application, CARE developers recommend permitting a maximum of two "does not apply" or missing responses per measure and disregarding any measure with over two responses of this kind during analysis. For up to two non-applicable or missing responses, they suggest scoring those responses with the average score for the remaining items of the measure, a mean-item score. Validation of the CARE measure in 2004 showed strong convergent high internal reliability (Cronbach's alpha, 0.93). In 2020, the CARE measure was adapted and validated for use in primary care consultations conducted in the Spanish language (Sp-CARE) [24], demonstrating high acceptability and face validity (1% of non-applicable or missing response), strong homogeneity (corrected item-total correlations >0.30), and robust internal reliability (Cronbach's alpha, 0.95).

2.3. Contextual adaptation of the Sp-CARE version

The adaptation of the Sp-CARE version to simulation followed a sequential process aimed at maintaining conceptual equivalence within this version and the new Sp-SIMCARE questionnaire. The process consisted of the following phases:

2.3.1. Phase 1: Preparation of a preliminary contextual adaptation proposal

In phase 1, two specialists in development and validation of competence evaluation scales prepared a proposal for preliminary contextual adaptation.

2.3.2. Phase 2: Sp-SIMCARE first version drafting

At this stage, a multi-disciplinary panel of five reviewers who had experience in conducting adaptation and psychometric validation of measurement tools, all of them professors at the faculty of medicine and health sciences at UIC, reviewed all items of the Sp-CARE scale and compared it to each item of the preliminary contextual adaptation proposal. The objective of the comparison was to assess the clarity, relevance, accuracy, and equivalence of meaning of the proposal.

Each item of the proposal was classified as either

1. conceptually equivalent and easily comprehensible,
2. functionally equivalent but with semantic discrepancies or comprehension difficulties, or
3. of uncertain equivalence.

If an item was classified as (2) or (3), the reviewers were required to clarify the rationale for the mismatch. The first version of Sp-SIMCARE questionnaire was consolidated after resolving the mismatches that had been identified.

2.3.3. Phase 3: Pilot evaluation of Sp-SIMCARE first version with SPs

In phase 3, five SPs assessed the first delivery of the Sp-SIMCARE questionnaire and provided responses to supplementary questions concerning comprehension of all items. These simulated patients also highlighted any irrelevant or offensive wording and were given the opportunity to suggest additional items for inclusion in the questionnaire. Their consultations with simulated patients, coupled with further refinements and improvements that they recommended, prompted the rewording of item 7, which was initially described inappropriately.

2.3.4. Phase 4: Review and refinement of the Sp-SIMCARE final version

During phase 4, the expert panel that had participated in phase 2 debated on the significance of the comments made by the SPs after the pilot evaluation until they reached a consensus, refining the final version of the SP-SIMCARE (see attachment 1, tables S1 and S2).

2.4. Validation of the Sp-SIMCARE final version

The Sp-SIMCARE questionnaire was validated through encounters of medical students with SPs. Four clinical scenarios, each portraying a common primary care situation, were designed for these encounters. In these scenarios, students interacted with SPs who played the roles of acute, chronic, functional, and patients with a hostile attitude (provoking situations difficult to deal with). To prevent the sharing of information among students, four different clinical cases were developed for each scenario. All clinical cases could be managed within a primary care setting without the necessity for referral to hospital care (see table 1 with detailed description of clinical scenarios and characteristics of SPs). Age and gender of undergraduates and patient simulation scenarios were assessed to identify differences in empathy between groups. Psychometric parameters that were evaluated included convergent validity, acceptability and face validity, homogeneity, and internal reliability, as outlined below:

2.4.1. Face validity

The scale's face validity was evaluated indirectly by the percentage of unanswered non-applicable responses and unanswered items in each survey. Up to two "does not apply" or missing responses were considered acceptable for each survey and substituted by the mean score for the remaining items in compliance with the criteria endorsed by the developers of the CARE measure.

2.4.2. Convergent validity

SPs scored undergraduates' level of empathy globally in response to the direct single question of "Is the student empathetic?" on a Likert scale of 1-10, with greater scores reflecting increased empathy. This global score was considered as the students' final grade in the subject. Therefore, we could consider it as the gold standard for assessing empathy for assessing any student's competency. Convergent validity of scores between the Sp-SIMCARE questionnaire and that global score of empathy was assessed by Pearson or Spearman correlations.

2.4.3. Homogeneity

Homogeneity was examined by corrected item-total correlations, where values above 0.30 predict high correlation [25].

2.4.4. Reliability

Cronbach's alpha was calculated to assess reliability and determine whether removal of any of the 10 items affected the consistency of the Sp-SIMCARE scale. Alpha values above 0.70 were considered satisfactory [25]. The processes of adaptation and validation of the new instrument are depicted in figure 1.

2.5. Statistical analysis

Normality of distribution of empathy scores was tested by the Shapiro-Wilk test. The Sp-SIMCARE scores were summarized as mean values and standard deviations or median and interquartile range values according to the normal or skewed distribution of data, respectively, and categorical variables were described as proportions. Comparison of scores by gender, age group, and simulated scenario was performed using the student t or the Mann-Whitney test (two-group comparison) and the ANOVA or the Kruskal-Wallis test (multi-group comparisons). The level of significance was set at 5% ($p < 0.05$). Data analysis was performed using the statistical package Stata v. 15. All the identifying information of the students was duly anonymized.

Table 1: Overview of clinical scenarios and simulated patient characteristics

Clinical scenario	Scenario brief description	Simulated patient characteristics
Common acute condition frequently seen in primary care	For example, urinary tract infection, low back pain, anxiety attack, etc. The clinical setting and the patient – who exhibits high levels of worry or fear – create a certain pressure on the student to process the medical history information and make clinical decisions.	<ul style="list-style-type: none"> ▪ Woman, 52 years old. Family medicine specialist (+5-year experience in clinical simulation) ▪ Man, 64 years old. Family medicine specialist (+10-year experience in clinical simulation)
Highly prevalent chronic condition in primary care	For example, type 2 diabetes mellitus, hypertension, chronic obstructive pulmonary disease (COPD), etc. This very common type of patient allows training of key competencies such as health education and treatment adjustment.	<ul style="list-style-type: none"> ▪ Woman, 69 years old. Nurse (+10-year experience in clinical simulation) ▪ Man, 70 years old. Physiotherapist (+5-year experience in clinical simulation)
“Functional” disorder (no identifiable organic cause); clinical interview conducted in the presence of a family member or other companion	Patient without organic pathology, whose psychosocial characteristics play a significant role. The companion interferes at times, complicating the doctor-patient relationship. The patient reports disabling symptoms (pain, fatigue, dizziness, etc.), but physical examination and complementary tests do not reveal an underlying cause, creating tension as the patient may perceive that the student, in their role as physician, doubts their suffering.	<ul style="list-style-type: none"> ▪ Woman, 67 years old. Family medicine specialist (+5-year experience in clinical simulation) ▪ Woman, 63 years old. Nurse (+ 10-year experience in clinical simulation)
Patient with a hostile attitude	Patient who generates discomfort, unease, and rejection. More than the possible pathology (usually a prevalent, non-acute condition), the patient’s negative attitude and hostility create immediate communication difficulties. The patient’s aggressiveness hinders an adequate medical history taking and/or explanation of treatments.	<ul style="list-style-type: none"> ▪ Man, 45 years old. Emergency medical assistant (+10-year experience in clinical simulation) ▪ Man, 42 years old. Emergency medicine specialist (+5-year experience in clinical simulation)

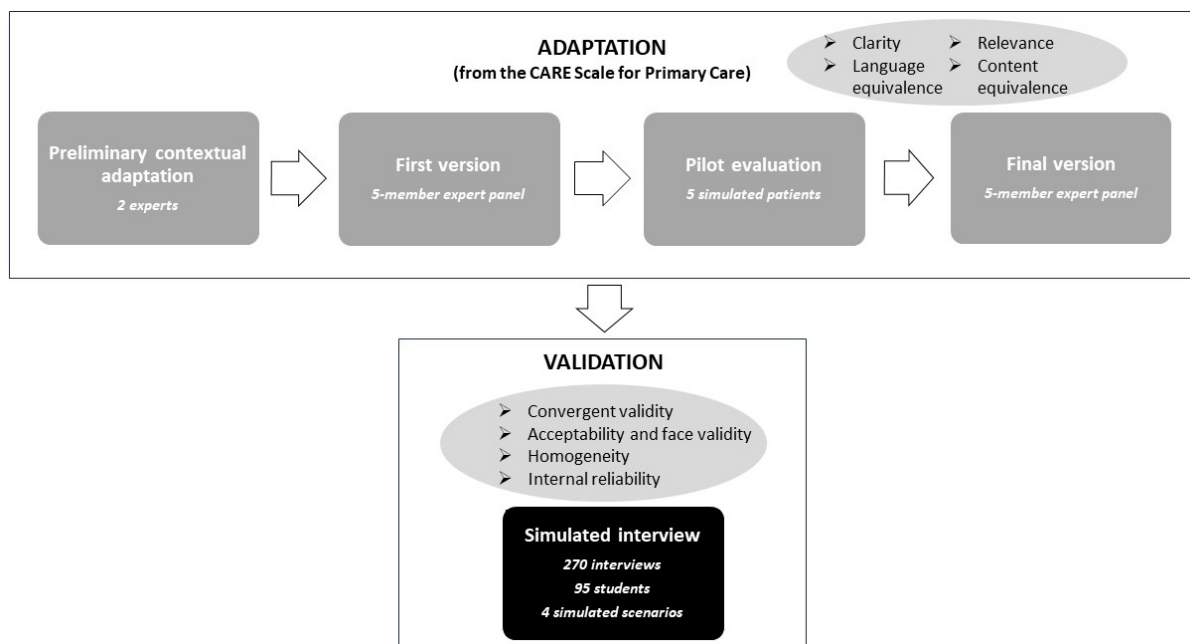


Figure 1: Diagram flow of the Sp-SIMCARE adaptation and validation process

Table 2: Medical students' demographics and Sp-SIMCARE scores by age group, gender, and simulated scenario

Variable	n (%)	Median score (IQR)	p
Age group			0.57
≤ 25 yr	79 (83.2)	32.5 (28.7-36.0)	
> 25 yr	16 (16.8)	33.0 (30.5-34.8)	
Gender			0.49
Male	35 (36.8)	32.4 (29.5-35.8)	
Female	60 (63.2)	33.5 (28.7-36.4)	
Simulation encounters			0.08
Acute patient	74 (27.4)	32.3 (28.0-38.0)	
Chronic patient	75 (27.8)	30.0 (23.0-39.0)	
Functional patient	74 (27.4)	34.0 (29.0-37.0)	
Hostile patient *	47 (17.4)	30.0 (23.0-39.0)	

* Significant score differences ($p \leq 0.001$) between this group and the rest of simulated patient groups

3. Results

The contextual adaptation process ensured the Sp-SIMCARE items were relevant, clearly worded, understandable, and had equivalent meaning for all users. A total of 95 students, 83% of whom were under 26 years old and 63.2% females, and eight SPs encountered in the four scenarios for the validation of the new instrument. Due to various contingencies, primarily related to students' absences, 297 out of the scheduled 380 encounters between students and SPs were completed (see table 2). Among the 297 SP evaluations of these encounters, 27 (9.1%) had three or more non-applicable or missing responses and were excluded from analysis, in line with the criteria recommended by the CARE developers.

The validation procedure eventually involved the examination of students' performance in 270 encounters that resulted in a total of 2,700 items of the Sp-SIMCARE be-

ing scored (see table 2). The median score was 32.5 (IQR, 29.0-37.0), with scores ranging from 10 to 50. Most of the responses to the individual items of the scale were scored as either "good" (45.2%) or "very good" (34.1%). There were no statistically significant in levels of empathy amongst the participants, based on age, gender or the type of simulated patient (see table 3).

The Sp-SIMCARE questionnaire demonstrated high face validity: only 53 (1.96%) out of the total 2,700 responses were "does not apply" choices or left blank (see table 2). Out of 270 surveys, there were 21 (7.8%) with two non-applicable or blank responses, and 11 (4.1%) with one of such responses. The encounters revealed that such responses were most frequent among difficult-to-deal-with patients (32/470, 6.81%), whereas the proportion was notably lower for chronic patients (19/750, 2.53%), and absent for acute or functional patients (see attachment 1, table S2). It was noted that a significant propor-

Table 3: Sp-SIMCARE scores for medical student performance

Item	Poor n (%)	Fair n (%)	Good n (%)	Very good n (%)	Excellent n (%)	Blank n (%)	Total n (%)
1. Making you feel at ease	9 (3.3)	25 (9.3)	118 (43.7)	104 (38.5)	14 (5.2)	0 (0.0)	270 (100.0)
2. Letting tell your 'story'	1 (0.4)	20 (7.4)	120 (44.4)	117 (43.3)	12 (4.4)	0 (0.0)	270 (100.0)
3. Really listening	4 (1.5)	22 (8.1)	108 (40.0)	125 (46.3)	11 (4.1)	0 (0.0)	270 (100.0)
4. Being interested in you as a whole person	5 (1.8)	35 (13.0)	119 (44.1)	102 (37.8)	9 (3.3)	0 (0.0)	270 (100.0)
5. Fully understanding your concerns	8 (3.0)	45 (16.7)	133 (49.3)	75 (27.8)	9 (3.3)	0 (0.0)	270 (100.0)
6. Showing care and compassion	7 (2.6)	38 (14.1)	130 (48.1)	87 (32.2)	8 (3.0)	0 (0.0)	270 (100.0)
7. Being positive	7 (2.6)	26 (9.6)	124 (45.9)	100 (37.0)	9 (3.3)	4 (1.5)	270 (100.0)
8. Explaining things clearly	5 (1.8)	40 (14.8)	128 (47.4)	87 (32.2)	10 (3.7)	3 (1.0)	270 (100.0)
9. Helping you to take control	7 (2.6)	49 (18.1)	127 (47.0)	54 (20.0)	7 (2.6)	26 (9.6)	270 (100.0)
10. Making a plan of action with you	9 (3.3)	46 (17.0)	114 (42.2)	69 (25.6)	9 (3.3)	23 (8.5)	270 (100.0)

Table 4: Convergent validity, homogeneity, and internal reliability of the Sp-SIMCARE questionnaire

Item	Convergent validity*	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha
1. Making you feel at ease	0.735	29.00	0.884	0.955
2. Letting tell your 'story'	0.590	28.89	0.827	0.957
3. Really listening	0.590	28.90	0.871	0.955
4. Being interested in you as a whole person	0.589	29.06	0.870	0.955
5. Fully understanding your concerns	0.607	29.21	0.848	0.957
6. Showing care and compassion	0.685	29.14	0.896	0.954
7. Being positive	0.644	29.05	0.850	0.956
8. Explaining things clearly	0.611	29.12	0.856	0.956
9. Helping you to take control	0.593	29.34	0.847	0.957
10. Making a plan of action with you	0.553	29.28	0.840	0.957

*All correlations statistically significant at a $p < 0.001$

tion of respondents either selected “does not apply” or left blank responses for items 9 “assist you in taking charge” (26/270, 9.6%) and 10 “collaborate with you to formulate a plan of action” (23/270, 8.5%). The proportion of non-applicable or missing responses to items 7 (4/270, 1.5%) and 8 (3/270, 1.0%) was much lower. Items 1 to 6 were scored in all surveys.

A significant positive correlation (Spearman's rho coefficient, 0.730; $p < 0.001$) was found between the scores provided by the simulated patients by the Sp-SIMCARE scale and in response to the query “Is the student empathetic?”, after confirming the skewed distribution of scores. Corrected item-total correlations ranged from 0.797 to 0.869, and Cronbach's alpha value for the scale was 0.960, with values for individual items falling within the range of 0.954-0.957 (see table 4).

4. Discussion

This study presents evidence of high validity and reliability of a new questionnaire, adapted from the well-known CARE instrument, which standardises the assessment of relational empathy in the context of clinical situations with simulated patients. The psychometric results support this tool as suitable, valid, and potentially useful for use with medical students in such settings.

Based on the results of the study, the new questionnaire demonstrated high acceptability with only 1.96% of “not applicable or blank” responses, which is comparable to the 1% observed in the validation of the previous Sp-CARE version [24]. This outcome indicates that the interactions generated in the simulation scenarios were realistic, varied, and effectively evaluated using the new questionnaire. Previous studies conducted in other primary care settings have documented a wide spectrum of acceptability of the original CARE measure and derived versions, with some works reporting similarly low rates of non-applicable or missing responses [16], [24] and others describing much higher percentages [14], [17]. The scenarios involving simulated patients difficult to deal with received the highest percentage of “not applicable or blank” re-

sponses (6.8%). These responses were concentrated in two specific items: 9 (9.6%) and 10 (8.5%). Those percentages were consistent with similar proportions of face validity for items 9 and 10 that were reported in earlier research on the CARE measure [12], [14], [15], [16]. Low face validity for the two items was predictable since management of emotions and patient containment prevailed over the expected performance of students when dealing with conflictive patients. In this regard, a previous study even postulated that item 10 should be excluded from scoring when assessing relational empathy, as it may not be an accurate determinant of a medical doctor's empathy but rather reflect shared decision making [26]. The median Sp-SIMCARE score obtained from our study population (32.5) was significantly lower than mean or median scores (above 40) previously published in CARE validation or implementation studies conducted in European primary care settings [12], [16]. The notable variation in scores between our and other European studies could be explained by the different study populations under evaluation: our study assessed the performance of undergraduate medical students who had limited prior experience interacting with simulated patients, while other studies evaluated the performance of primary care medical doctors who commonly have regular and intense interactions with their patients and, as a result, are more likely to exhibit empathetic competency. Interestingly, no significant variations were found in Sp-SIMCARE scores based on the gender of undergraduate medical students. This outcome aligns with prior studies, which suggested that the scores by the CARE measure were not substantially affected by either the gender of medical professionals or consultation characteristics [13], [14], [15], [24]. Furthermore, although the population studied here (fourth-year students) may have had an initial imbalance in empathy in favour of women on entering medical school, this difference could have been compensated for by the training received throughout the years of study, as we showed in [27].

The Sp-SIMCARE questionnaire displayed robust convergence (Spearman's rho coefficient, 0.730) with the scores

provided by simulated patients in response to the explicit query “Is the student empathetic?”, which were used for global evaluation of the simulation exercise. This outcome indicates that the novel scale measures students’ empathy levels in a manner that aligns with the comprehensive assessment of empathy conducted by simulated patients in the four most common types of clinical patients (chronic, acute, functional, and difficult-to-deal with), giving validity to the use of the Sp-SIMCARE scale in different simulation contexts. Overall, corrected item-total correlation values (>0.797) and Cronbach’s alpha values (>0.954) were high in our study and revealed strong homogeneity and reliability of the new tool, in line with values reported for these measures in other previous validation studies of CARE versions [13], [14], [15], [24].

The present work logically presents some methodological strengths and limitations. Among the strengths, it is worth highlighting the meticulous sequential process followed to adapt Sp-CARE to the simulation context. Furthermore, the fact that it was tested in different pathology scenarios and by simulated patients with different profiles suggests that the tool has a good usability. In this sense, the scenarios, which had undergone a prior design and validation process by a committee of experts, considered not only different clinical situations, but also different types of patients, including a specific scenario with hostile SPs characterized by lack of cooperation or high aggressiveness. One limitation of the study is the absence of an objective gold-standard of relational empathy. Therefore, to assess the convergent validity of the Sp-SIMCARE questionnaire, it was necessary to compare its results with a proxy for the gold standard. In this case, we did not use a validated survey to measure empathy but instead relied on the simulated patient’s self-reported perception of empathy during the clinical encounter. This served as our external standard or criterion. While some may consider this a limitation, it is a commonly used procedure in similar cases. On the other hand, the study population size was small, which restricted the possibility of exploratory and confirmatory factor analyses. Additionally, the origin of the study population, consisting of fourth-year student volunteers, suggests the need for caution when generalizing the results.

The Sp-SIMCARE questionnaire’s focus on the relational aspects of empathy provides an advantage. It assesses how empathy translates into concrete actions during simulated interactions, offering valuable insight beyond a subjective and emotional understanding. The questionnaire is particularly sensitive to the complexities of interacting with varied scenarios and patients, including the most difficult ones. Adaptability is crucial in medical training, particularly when testing empathy in challenging clinical situations, such as with conflictive patients. To be highlighted, the use of simulated environments for questionnaire validation is a strategic choice. It provides a controlled and safe environment, ensuring a consistent and fair assessment for all students, especially those who are still learning. This methodology addresses the practical barriers associated with obtaining direct feed-

back from patients by providing a structured and reproducible assessment. Importantly, the validation of Sp-SIMCARE in advanced medical students suggests its potential usefulness for assessing relational empathy in other health care professional groups, although further studies are needed to confirm this.

5. Conclusion

In summary, the Sp-SIMCARE questionnaire proved to be psychometrically valid and reliable for evaluation of undergraduate medical students by simulated patients. The questionnaire’s uniqueness lies in its ability to measure the relational dimension of empathy, providing a practical tool for assessing this competence. The use of this new tool could potentially assist in the design and implementation of interventions aimed at fostering empathy in future doctors throughout their training.

Acknowledgements

The authors acknowledge the contribution of simulated patients and students enrolled in the Degree in Medicine of the Universitat Internacional de Catalunya that participated in this study.

Notes

Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee at the study centre (registration MED-2022-07). The approval of the protocol waived the requirement to obtain informed consent from participants in the study.

Funding

This study was partially funded by a competitive grant for medical education research projects awarded by the Sociedad Española de Educación Médica (SEDEM).

Authorship

Pedro Brotons de los Reyes and Montserrat Virumbrales Cancio share the first authorship.

Author contributions

- PB: Data analysis, interpretation of results, draft manuscript preparation, final manuscript approval
- MV: Study conception and design, data collection, interpretation of results, final manuscript approval
- PC: Interpretation of results, final manuscript approval
- XM: Data collection, interpretation of results, final manuscript approval
- SD: Data collection, interpretation of results, final manuscript approval
- AB: Study conception and design, study management, interpretation of results, final manuscript approval

Authors' ORCIDs

- Pedro Brotos de los Reyes: [0000-0002-2399-7320]
- Montserrat Virumbrales Cancio: [0000-0003-3541-2948]
- Pere Castellvi: [0000-0002-3920-8576]
- Xavier Martinez-Regada: [0009-0009-5013-7738]
- Albert Balaguer: [0000-0002-5222-8635]

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from <https://doi.org/10.3205/zma001844>

1. Attachment_1.pdf (303 KB)
Supplementary material

References

1. Derksen F, Olde Hartman TC, van Dijk A, Plouvier A, Bensing J, Lagro-Janssen A. Consequences of the presence and absence of empathy during consultations in primary care: a focus group study with patients. *Patient Educ Couns*. 2017;100(5):987-993. DOI: 10.1016/j.pec.2016.12.003
2. Hojat M, LaNoue M. Exploration and confirmation of the latent variable structure of the Jefferson scale of empathy. *Int J Med Educ*. 2014;20(5):73-81. DOI: 10.5116/ijme.533f.0c41
3. Mercer SW, Reynolds WJ. Empathy and quality of care. *Br J Gen Pract*. 2002;52 Suppl(Suppl):S9-S12.
4. Hall JA, Schwartz R, Duong F, Niu Y, Dubey M, DeSteno D, Sanders JJ. What is clinical empathy? Perspectives of community members, university students, cancer patients, and physicians. *Patient Educ Couns*. 2021;104(5):1237-1245. DOI: 10.1016/j.pec.2020.11.001
5. Schwartz R, Dubey M, Blanch-Hartigan D, Sanders JJ, Hall JA. Physician empathy according to physicians: A multi-specialty qualitative analysis. *Patient Educ Couns*. 2021;104(10):2425-2431. DOI: 10.1016/j.pec.2021.07.024
6. Sanders JJ, Dubey M, Hall JA, Catzen HZ, Blanch-Hartigan D, Schwartz R. What is empathy? Oncology patient perspectives on empathic clinician behaviors. *Cancer*. 2021;127(22):4258-4265. DOI: 10.1002/cncr.33834
7. Decety J, Svetlova M. Putting together phylogenetic and ontogenetic perspectives on empathy. *Dev Cogn Neurosci*. 2012;2(1):1-24. DOI: 10.1016/j.dcn.2011.05.003
8. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Heal Prof*. 2022;27(3):237-251. DOI: 10.1177/0163278704267037
9. Del Canale S, Louis DZ, Maio V, Wang X, Rossi G, Hojat M, Gonnella JS. The relationship between physician empathy and disease complications: an empirical study of primary care physicians and their diabetic patients in Parma, Italy. *Acad Med*. 2012;87(9):1243-1249. DOI: 10.1097/ACM.0b013e3182628fbf
10. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' empathy and clinical outcomes for diabetic patients. *Acad Med*. 2011;86(3):359-364. DOI: 10.1097/ACM.0b013e3182086fe1
11. Hemmerdinger JM, Stoddart SD, Lilford RJ. A systematic review of tests of empathy in medicine. *BMC Med Educ*. 2007;7:24. DOI: 10.1186/1472-6920-7-24
12. Mercer SW, Maxwell M, Heaney D, Watt GC. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract*. 2004;21(6):699-705.
13. Mercer SW, McConnachie A, Maxwell M, Heaney D, Watt GC. Relevance and practical use of the Consultation and Relational Empathy (CARE) Measure in general practice. *Fam Pract*. 2005;22(3):328-334. DOI: 10.1093/fampra/cm621
14. Fung CS, Hua A, Tam L, Mercer SW. Reliability and validity of the Chinese version of the CARE measure in a primary care setting in Hong Kong. *Fam Pract*. 2009;26(5):398-406. DOI: 10.1093/fampra/cmp044
15. Aomatsu M, Abe H, Abe K, Yasui H, Suzuki T, Sato J, Ban N, Mercer SW. Validity and reliability of the Japanese version of the CARE measure in a general medicine outpatient setting. *Fam Pract*. 2014;31(1):118-126. DOI: 10.1093/fampra/cmt053
16. van Dijk I, Scholten Meilink Lenferink N, Lucassen PL, Mercer SW, van Weel C, Hartman TC, Speckens AE. Reliability and validity of the Dutch version of the Consultation and Relational Empathy Measure in primary care. *Fam Pract*. 2017;34(1):119-124. DOI: 10.1093/fampra/cmw116
17. Hanževački M, Jakovina T, Bajić Ž, Tomac A, Mercer S. Reliability and validity of the Croatian version of Consultation and Relational Empathy (CARE) Measure in primary care setting. *Croat Med J*. 2015;56(1):50-56. DOI: 10.3325/cmj.2015.56.50
18. Solaja O, Moloo H, Hopkins E, Khan N, Gilbert RW, Neville A, Kubelik D, Maziak D, Rowe N, Odell M, Momtazi M. Implementation, results and face validity of the Consultation and Relational Empathy measure in a Canadian department of surgery. *Can J Surg*. 2022;65(6):E749-E755. DOI: 10.1503/cjs.003721
19. Del Barrio LG, Rodríguez-Díez C, Gea A, Arbea L, Pereira J, Díez N. Impact of a longitudinal course on medical professionalism on the empathy of medical students. *Patient Educ Couns*. 2024;119:108042. DOI: 10.1016/j.pec.2023.108042
20. Duran A, Donelan C, Bowman Peterson J, Gladding SP, Weissmann P, Roth CS. Communicating value to patients—a high-value care communication skills curriculum. *Postgrad Med*. 2021;133(2):231-236. DOI: 10.1080/00325481.2020.1807728
21. Schreckenbach T, Ochsendorf F, Sterz J, Rüsseler M, Bechstein WO, Bender B, Bechtoldt MN. Emotion recognition and extraversion of medical students interact to predict their empathic communication perceived by simulated patients. *BMC Med Educ*. 2018;18(1):237. DOI: 10.1186/s12909-018-1342-8

22. Vogel D, Meyer M, Harendza S. Verbal and non-verbal communication skills including empathy during history taking of undergraduate medical students. *BMC Med Educ.* 2018;18(1):157. DOI: 10.1186/s12909-018-1260-9
23. Buffel du Vaure C, Lemogne C, Bunge L, Catu-Pinault A, Hoertel N, Ghasarossian C, Vincens ME, Galam E, Jaury P. Promoting empathy among medical students: A two-site randomized controlled study. *J Psychosom Res.* 2017;103:102-107. DOI: 10.1016/j.jpsychores.2017.10.008
24. García Del Barrio L, Rodríguez-Díez C, Martín-Lanas R, Costa P, Costa MJ, Díez N. Reliability and validity of the Spanish (Spain) version of the consultation and relational empathy measure in primary care. *Fam Pract.* 2021;38(3):353-359. DOI: 10.1093/fampra/cmab135
25. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. New York (NY): Oxford University Press; 2008.
26. Wirtz M, Boecker M, Forkmann T, Neumann M. Evaluation of the "Consultation and Relational Empathy" (CARE) measure by means of Rasch-analysis at the example of cancer patients. *Patient Educ Couns.* 2011;82(3):298-306. DOI: 10.1016/j.pec.2010.12.009
27. Brotos P, Virumbrales M, Elorduy M, Díaz de Castellví S, Mezquita P, Gené E, Balaguer A. Improvement of medical students' performance in simulated patient interviews by pre-clinical communication training. *Int J Med Educ.* 2022;13:148-153. DOI: 10.5116/ijme.6299.c15f

Corresponding author:

Albert Balaguer
 Universitat Internacional de Catalunya, School of Medicine and Health Sciences, Department of Medicine, Josep Trueta Street, s/n. ES-08195 Sant Cugat del Vallès, Spain
 abalaguer@uic.es

Please cite as

Brotos de los Reyes P, Virumbrales Cancio M, Castellví P, Martínez-Regada X, Balaguer A. Assessment of empathy by simulated patients: Adaptation and validation of a new instrument. *GMS J Med Educ.* 2026;43(4):Doc50. DOI: 10.3205/zma001844, URN: urn:nbn:de:0183-zma0018446

This article is freely available from

<https://doi.org/10.3205/zma001844>

Received: 2024-10-03

Revised: 2025-06-08

Accepted: 2025-10-27

Published: 2026-04-15

Copyright

©2026 Brotos de los Reyes et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Bewertung der Empathie durch simulierte Patienten: Anpassung und Validierung eines neuen Instruments

Zusammenfassung

Hintergrund: Das Einfühlungsvermögen von Ärzt:innen ist für deren Interaktion mit Patient:innen von entscheidender Bedeutung und wirkt sich auf das Vertrauen der Patient:innen und die Gesundheitsergebnisse aus. Die CARE-Skala (Consultation and Relational Empathy) hilft bei der Bewertung des Einfühlungsvermögens von medizinischem Fachpersonal, jedoch bestehen begründete Zweifel an ihrer Validität, wenn sie direkt in Simulationskontexten eingesetzt wird. Diese Studie adaptiert und validiert eine Version der CARE-Skala für simulierte Patienten (Sp-SIMCARE) und schließt damit eine Lücke in der Einfühlungsvermögensbewertung innerhalb medizinischer Trainingssimulationen.

Methoden: Die kontextuelle Anpassung der CARE-Skala für simulierte Patient:innen erfolgte in vier Phasen: 1) Vorbereitung eines vorläufigen Anpassungsvorschlags; 2) Entwurf der ersten Version; 3) Erprobung der ersten Version mit simulierten Patient:innen; und 4) Überarbeitung der endgültigen Sp-SIMCARE-Version. Ein Gremium aus fünf Expert:innen arbeitete mit fünf simulierten Patient:innen zusammen, um Klarheit, Relevanz und sprachliche Äquivalenz sicherzustellen. Die Validierung erfolgte anhand typischer Szenarien aus der Primärversorgung. Simulierte Patient:innen bewerteten die Leistung von Medizinstudierenden in vier primären Patient:innenrollen: akut, chronisch, mit hoher funktionseller Komponente und mit einer feindseligen Haltung. Zu den bewerteten psychometrischen Parametern gehörten die konvergente Validität (bewertet durch simulierte Patient:innen anhand einer Gesamtpunktzahl), die Akzeptanz und Augenscheinvalidität, die Homogenität und die interne Reliabilität.

Ergebnisse: Der Anpassungsprozess führte zu einer klaren, relevanten und verständlichen Skala, die ein einheitliches Verständnis unter den Anwendern gewährleistet. An der Validierung nahmen 95 Studierende in 270 Begegnungen mit acht simulierten Patient:innen teil. Die endgültige Version zeigte konvergente Validität (Spearman's $\rho=0,730$, $p<0,001$), Akzeptanz und Augenscheinvalidität (1,96% „trifft nicht zu“/fehlende Werte), Homogenität (Item-Gesamt-Korrelationen 0,705–0,865) und Reliabilität (Cronbach's $\alpha=0,960$).

Schlussfolgerungen: Die Sp-SIMCARE-Skala scheint ein valides und zuverlässiges Instrument für simulierte Patient:innen zu sein, um das Einfühlungsvermögen angehender Ärzt:innen während ihrer Ausbildung aus einer multidimensionalen Perspektive zu bewerten.

Schlüsselwörter: Einfühlungsvermögen, medizinische Ausbildung, Medizinstudent, Patientensimulation, Validierungsstudie, Psychometrie

1. Einleitung

Die Fähigkeit von Ärzt*innen, sich in ihre Patient*innen hineinzuversetzen, ist für den Aufbau einer vertrauensvollen und effektiven Arzt-Patient*innen-Beziehung unerlässlich und führt zu besseren klinischen Ergebnissen [1]. Trotz der Betonung ihrer Bedeutung gibt es keinen Konsens über die Definition von Empathie im spezifischen

klinischen Kontext. Im Allgemeinen wird sie als kognitive Fähigkeit beschrieben, die inneren Erfahrungen und Perspektiven eines Patient*innen zu verstehen und dieses Verständnis zu kommunizieren [2]. Einige Autor*innen unterscheiden vier Schlüsselkomponenten der Empathie: emotionale (die Fähigkeit, den emotionalen Zustand des Patient*in zu teilen), moralische (die Motivation, empathisch zu agieren), kognitive (die Fähigkeit, die Perspektive des Patient*in zu verstehen) und verhaltensbezogene (die Fähigkeit, mit dem Patient*in auf hilfreiche oder

Pedro Brotons de los Reyes^{1,2,3}
Montserrat Virumbrales Cancio¹
Pere Castellvi¹
Xavier Martínez-Regada^{1,4}
Albert Balaguer^{1,2,5}

- 1 Internationale Universitat de Catalunya, Fakultät für Medizin und Gesundheitswissenschaften, Abteilung für Medizin, Sant Cugat del Vallès, Spanien
- 2 Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spanien
- 3 Instituto de Salud Carlos III, Consorcio de Investigación Biomédica en Red de Epidemiología y Salud (CIBERESP), Madrid, Spanien
- 4 EAP Amadeu Torner, L'Hospitalet de Llobregat, Barcelona, Spanien
- 5 Hospital Universitari General de Catalunya, Department of Pediatrics, Barcelona, Spain

therapeutische Weise zu kommunizieren und zu handeln) [3]. Trotz unterschiedlicher Modelle sind sich die meisten Forscher*innen in den letzten Jahren einig geworden, dass es für Kliniker wichtig ist, die emotionalen Zustände ihrer Patient*innen zu erkennen und durch Verhaltensreaktionen zum Ausdruck zu bringen [4], [5], [6].

Die soziale Neurowissenschaft hat in letzter Zeit Fortschritte dabei gemacht, Empathiekomponenten zu identifizieren, die durch Veränderungen in den Erfahrungen modifiziert werden können [7]. Daher ist Empathie eine Kompetenz, die gelehrt und gelernt werden kann, insbesondere in den frühen Phasen der medizinischen Praxis. Eine Schulung in dieser Kompetenz könnte die Fähigkeit von medizinischem Fachpersonal und Studierenden verbessern, eine wirksame medizinische Versorgung zu leisten, und gleichzeitig die Erfahrungen und das Engagement der Patient*innen verbessern [8], [9], [10]. Zur Bewertung der Empathie von Ärzt*innen im Rahmen der Allgemeinmedizin stehen mehrere Instrumente zur Verfügung, darunter der Davis' Interpersonal Reactivity Index (IRI), die Jefferson Scale of Physician Empathy (JSPE) und die Consultation and Relational Empathy (CARE)-Skala [11]. Insbesondere bewertet die CARE-Skala

1. die Fähigkeit, die Situation, Perspektive und Gefühle des Patient*innen (und deren damit verbundene Bedeutungen) zu verstehen;
2. dieses Verständnis zu kommunizieren und dessen Richtigkeit zu überprüfen; und
3. auf der Grundlage dieses Verständnisses auf hilfreiche (therapeutische) Weise mit dem Patient*innen umzugehen [12].

Obwohl der IRI, der JSPE und andere analoge Instrumente relevante Merkmale für die Bewertung der emotionalen und kognitiven Dimensionen von Empathie aufweisen, zeichnet sich die CARE-Skala dadurch aus, dass sie auch die nachfolgende Verhaltensreaktion bewertet, die sowohl für die Patient*innenversorgung als auch für die Ausbildung von Medizinstudierenden und Fachkräften von größter Bedeutung ist [13].

Die CARE-Skala wurde in Primär- und Facharztpraxen in ganz Europa, Nordamerika und Asien übersetzt und validiert [14], [15], [16], [17], [18]. Darüber hinaus wurde sie direkt zur Bewertung von Empathie in Simulationskontexten eingesetzt [19], [20], [21], [22], [23]. Obwohl es naheliegend erscheint, dass ein Instrument, das im realen Kontext gute metrische Eigenschaften aufweist, auch in einem simulierten Kontext ähnliche und angemessene Ergebnisse liefert, sollte diese Hypothese getestet werden, um ihre Gültigkeit sicherzustellen.

Das Ziel dieser Studie war es, die spanische Version des CARE-Instruments (Sp-CARE) [24] für den Einsatz bei Medizinstudenten in simulierten Umgebungen anzupassen und seine metrischen Eigenschaften zu bewerten.

2. Methoden

2.1. Design und Studienpopulation

Diese Studie wurde an der Fakultät für Medizin und Gesundheitswissenschaften der Universitat Internacional de Catalunya (UIC) in Barcelona, Spanien, während des akademischen Jahres 2022-2023 durchgeführt. Zu den Teilnehmenden gehörten Medizinstudierende im vierten Studienjahr und SPs, die über umfangreiche Erfahrung in der Bewertung von „klinischen Kernkompetenzen“ wie zwischenmenschlichen Fähigkeiten, Kommunikations- und Zuhörfähigkeiten, Zeitmanagement, Problemlösung, Führungsqualitäten und Empathie verfügten und standardisierte Instrumente zur Messung von Empathie einsetzen. Das Studienprotokoll wurde von der Ethikkommission des Studienzentrums genehmigt (Registrierung MED-2022-07).

2.2. Das Messinstrument für Konsultations- und Beziehungsempathie

Das CARE-Maß wurde 2004 entwickelt und validiert und zeigt eine starke Korrelation mit anderen empathiebezogenen Skalen sowie eine hohe interne Reliabilität (Cronbachs Alpha, 0,93) [12]. Die Skala besteht aus 10 Items, die auf einer Likert-Skala von „1“ (schlecht) bis „5“ (ausgezeichnet) bewertet werden, wobei höhere Werte ein höheres Maß an Empathie innerhalb eines Bereichs von 10 bis 50 bedeuten. Die Entwickler der Messung schlagen vor, dass die Items 1 bis 6 in erster Linie mit dem affektiven Aspekt der Empathie zusammenhängen, während die übrigen Items mit kognitiven und verhaltensbezogenen Aspekten zusammenhängen. Für jeden Punkt gibt es auch die Option „trifft nicht zu“, die ausgewählt werden kann, wenn der Punkt für die Konsultation als irrelevant erachtet wird. Eine minimale Anzahl von „trifft nicht zu“-Antworten oder fehlenden Antworten zeigt somit die vom Patient*innen wahrgenommene Relevanz des Punktes. Um die praktische Anwendbarkeit zu gewährleisten, empfehlen die Entwickler von CARE, maximal zwei „trifft nicht zu“-Antworten oder fehlende Antworten pro Messung zuzulassen und alle Messungen mit mehr als zwei Antworten dieser Art bei der Analyse zu ignorieren. Bei bis zu zwei nicht zutreffenden oder fehlenden Antworten schlagen sie vor, diese Antworten mit der Durchschnittspunktzahl der übrigen Items der Maßnahme, also einer mittleren Item-Punktzahl, zu bewerten. Die Validierung der CARE-Maßnahme im Jahr 2004 ergab eine starke konvergente hohe interne Reliabilität (Cronbachs Alpha, 0,93). Im Jahr 2020 wurde die CARE-Maßnahme für den Einsatz in spanischsprachigen Hausarztpraxen angepasst und validiert (Sp-CARE) [24] und zeigte eine hohe Akzeptanz und Augenscheinvalidität (1% nicht zutreffende oder fehlende Antworten), eine starke Homogenität (korrigierte Item-Gesamt-Korrelationen > 0,30) und eine robuste interne Reliabilität (Cronbachs Alpha, 0,95).

2.3. Kontextuelle Anpassung der Sp-CARE-Version

Die Anpassung der Sp-CARE-Version an die Simulation erfolgte in einem sequenziellen Prozess, der darauf abzielte, die konzeptionelle Äquivalenz zwischen dieser Version und dem neuen Sp-SIMCARE-Fragebogen zu wahren. Der Prozess umfasste die folgenden Phasen:

2.3.1. Phase 1: Erstellung eines vorläufigen Vorschlags zur kontextuellen Anpassung

In Phase 1 erstellten zwei Spezialist*innen für die Entwicklung und Validierung von Kompetenzbewertungsskalen einen Vorschlag für eine vorläufige kontextuelle Anpassung.

2.3.2. Phase 2: Entwurf der ersten Version von Sp-SIMCARE

In dieser Phase überprüfte ein multidisziplinäres Gremium aus fünf Gutachter*innen alle Items der Sp-CARE-Skala und verglich sie mit jedem Item des vorläufigen Vorschlags zur kontextuellen Anpassung. Alle fünf Gutachter*innen hatten Erfahrung in der Anpassung und psychometrischen Validierung von Messinstrumenten und waren Professor*innen an der Fakultät für Medizin und Gesundheitswissenschaften der UIC. Ziel des Vergleichs war es, die Klarheit, Relevanz, Genauigkeit und Bedeutungsäquivalenz des Vorschlags zu bewerten. Jeder Punkt des Vorschlags wurde entweder als

1. konzeptionell äquivalent und leicht verständlich,
2. funktional äquivalent, aber mit semantischen Diskrepanzen oder Verständnisschwierigkeiten oder
3. von ungewisser Äquivalenz eingestuft.

Wenn ein Punkt als (2) oder (3) eingestuft wurde, mussten die Prüfer*innen die Gründe für die Nichtübereinstimmung darlegen. Die erste Version des Sp-SIMCARE-Fragebogens wurde nach Behebung der festgestellten Nichtübereinstimmungen konsolidiert.

2.3.3. Phase 3: Pilotbewertung der ersten Version von Sp-SIMCARE mit SPs

In Phase 3 bewerteten fünf SPs die erste Fassung des Sp-SIMCARE-Fragebogens und beantworteten ergänzende Fragen zum Verständnis aller Punkte. Diese simulierten Patient*innen wiesen auch auf irrelevante oder beleidigende Formulierungen hin und erhielten die Möglichkeit, zusätzliche Punkte für die Aufnahme in den Fragebogen vorzuschlagen. Ihre Konsultationen mit simulierten Patient*innen sowie die von ihnen empfohlenen weiteren Verfeinerungen und Verbesserungen führten zu einer Umformulierung von Punkt 7, der ursprünglich unangemessen beschrieben war.

2.3.4. Phase 4: Überprüfung und Verfeinerung der endgültigen Version von Sp-SIMCARE

In Phase 4 diskutierte das Expertengremium, das an Phase 2 teilgenommen hatte, über die Bedeutung der Kommentare der SPs nach der Pilotbewertung, bis es zu einem Konsens gelangte, und verfeinerte die endgültige Version von Sp-SIMCARE (siehe Anhang 1, Tabellen S1 und S2).

2.4. Validierung der endgültigen Version von Sp-SIMCARE

Der Sp-SIMCARE-Fragebogen wurde durch Begegnungen von Medizinstudierenden mit SPs validiert. Für diese Begegnungen wurden vier klinische Szenarien entworfen, die jeweils eine gängige Situation in der Primärversorgung darstellen. In diesen Szenarien interagierten die Studierenden mit SPs, die die Rollen von akuten, chronischen, funktionellen und Patienten mit einer feindseligen Haltung (die schwierige Situationen provozieren) spielten. Um den Austausch von Informationen zwischen den Studierenden zu verhindern, wurden für jedes Szenario vier verschiedene klinische Fälle entwickelt. Alle klinischen Fälle konnten innerhalb der Primärversorgung behandelt werden, ohne dass eine Überweisung an ein Krankenhaus erforderlich wäre (siehe Tabelle 1 mit einer detaillierten Beschreibung der klinischen Szenarien und Merkmale der SPs). Das Alter und Geschlecht der Studierenden sowie die Patient*innensimulationsszenarien wurden bewertet, um Unterschiede in der Empathie zwischen den Gruppen zu ermitteln. Zu den bewerteten psychometrischen Parametern gehörten die konvergente Validität, die Akzeptanz und die Augenscheinvalidität, die Homogenität und die interne Reliabilität, wie unten beschrieben:

2.4.1. Gesichtvalidität

Die Gesichtvalidität der Skala wurde indirekt anhand des Prozentsatzes der unbeantworteten, nicht zutreffenden Antworten und unbeantworteten Fragen in jeder Umfrage bewertet. Bis zu zwei „trifft nicht zu“-Antworten oder fehlende Antworten wurden für jede Umfrage als akzeptabel angesehen und gemäß den von den Entwicklern der CARE-Messung festgelegten Kriterien durch den Durchschnittswert der übrigen Fragen ersetzt.

2.4.2. Konvergente Validität

Die SPs bewerteten das Empathieniveau der Studierenden insgesamt anhand der direkten Frage „Ist der Studierende empathisch?“ auf einer Likert-Skala von 1 bis 10, wobei höhere Werte ein höheres Maß an Empathie widerspiegeln. Diese Gesamtpunktzahl wurde als Abschlussnote der Studierenden in diesem Fach gewertet. Daher konnten wir sie als Goldstandard für die Bewertung der Empathie zur Beurteilung der Kompetenz eines jeden Studierenden betrachten. Die konvergente Validität der Werte zwischen dem Sp-SIMCARE-Fragebogen und dieser

Tabelle 1: Übersicht über klinische Szenarien und simulierte Patientenmerkmale

Klinisches Szenario	Kurzbeschreibung des Szenarios	Simulierte Patientenmerkmale
<p>Häufige akute Erkrankung, die häufig in der Primärversorgung auftritt</p>	<p>Beispiele hierfür sind Harnwegsinfektionen, Schmerzen im unteren Rückenbereich, Angstattacken usw. Das klinische Umfeld und der Patient – der ein hohes Maß an Besorgnis oder Angst zeigt – üben einen gewissen Druck auf den Studenten aus, die Informationen zur Krankengeschichte zu verarbeiten und klinische Entscheidungen zu treffen.</p>	<ul style="list-style-type: none"> ▪ Frau, 52 Jahre alt. Fachärztin für Familienmedizin (+5 Jahre Erfahrung in der klinischen Simulation) ▪ Mann, 64 Jahre alt. Facharzt für Familienmedizin (+10 Jahre Erfahrung in der klinischen Simulation)
<p>Häufig auftretende chronische Erkrankung in der Primärversorgung</p>	<p>Beispiele hierfür sind Typ-2-Diabetes mellitus, Bluthochdruck, chronisch obstruktive Lungenerkrankung (COPD) usw. Diese sehr häufige Patientengruppe ermöglicht das Training wichtiger Kompetenzen wie Gesundheitsaufklärung und Behandlungsanpassung.</p>	<ul style="list-style-type: none"> ▪ Frau, 69 Jahre alt. Krankenschwester (+10 Jahre Erfahrung in der klinischen Simulation) ▪ Mann, 70 Jahre alt. Physiotherapeut (+5 Jahre Erfahrung in der klinischen Simulation)
<p>„Funktionale“ Störung (keine erkennbare organische Ursache); klinisches Gespräch in Anwesenheit eines Familienmitglieds oder einer anderen Begleitperson</p>	<p>Patient ohne organische Pathologie, dessen psychosoziale Merkmale eine bedeutende Rolle spielen. Der Begleiter mischt sich gelegentlich ein und erschwert dadurch die Arzt-Patient-Beziehung. Der Patient berichtet über behindernde Symptome (Schmerzen, Müdigkeit, Schwindel usw.), aber die körperliche Untersuchung und ergänzende Tests ergeben keine zugrunde liegende Ursache, was zu Spannungen führt, da der Patient möglicherweise den Eindruck hat, dass der Student in seiner Rolle als Arzt sein Leiden anzweifelt.</p>	<ul style="list-style-type: none"> ▪ Frau, 67 Jahre alt. Fachärztin für Familienmedizin (+5 Jahre Erfahrung in der klinischen Simulation) ▪ Frau, 63 Jahre alt. Krankenschwester (+10 Jahre Erfahrung in der klinischen Simulation)
<p>Patient mit einer feindseligen Haltung</p>	<p>Patient, der Unbehagen, Unruhe und Ablehnung hervorruft. Mehr noch als die mögliche Erkrankung (in der Regel eine weit verbreitete, nicht akute Erkrankung) führen die negative Einstellung und Feindseligkeit des Patienten zu unmittelbaren Kommunikationsschwierigkeiten. Die Aggressivität des Patienten behindert eine angemessene Anamnese und/oder die Erläuterung der Behandlungen.</p>	<ul style="list-style-type: none"> ▪ Mann, 45 Jahre alt. Rettungssanitäter (+10 Jahre Erfahrung in der klinischen Simulation) ▪ Mann, 42 Jahre alt. Facharzt für Notfallmedizin (+5 Jahre Erfahrung in der klinischen Simulation)

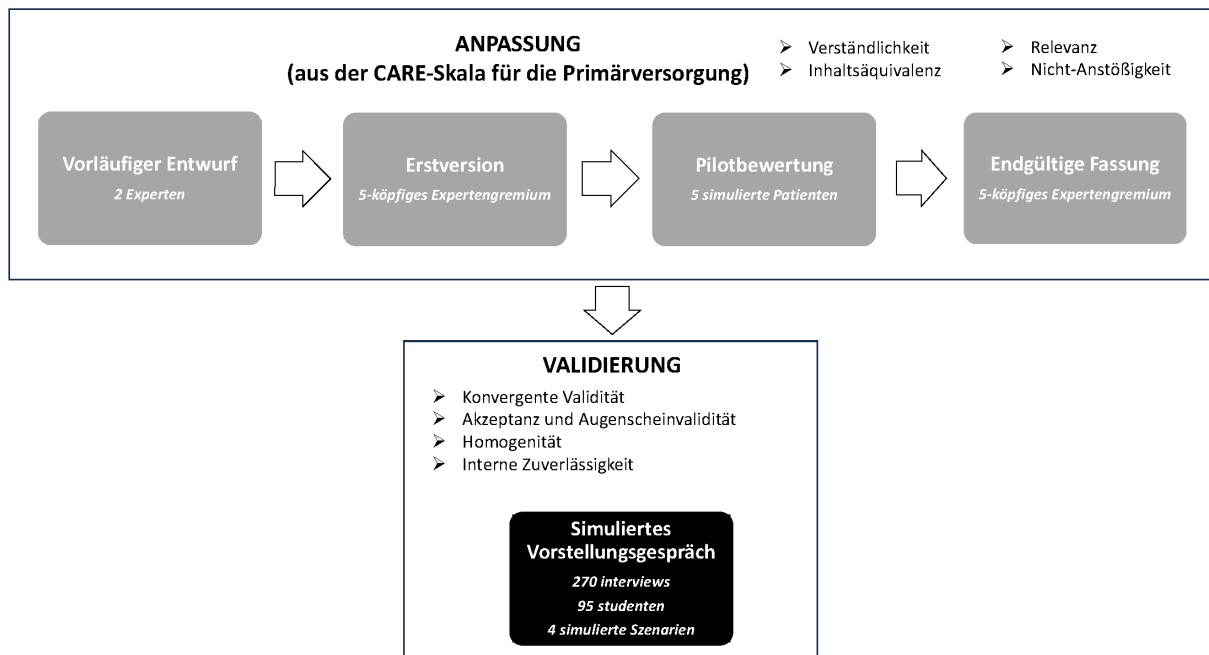


Abbildung 1: Diagramm zum Ablauf des Anpassungs- und Validierungsprozesses von Sp-SIMCARE

Gesamtbewertung der Empathie wurde anhand von Pearson- oder Spearman-Korrelationen bewertet.

2.4.3. Homogenität

Die Homogenität wurde anhand korrigierter Item-Gesamtkorrelationen untersucht, wobei Werte über 0,30 eine hohe Korrelation vorhersagen [25].

2.4.4. Zuverlässigkeit

Cronbachs Alpha wurde berechnet, um die Zuverlässigkeit zu bewerten und festzustellen, ob die Entfernung eines der 10 Items die Konsistenz der Sp-SIMCARE-Skala beeinträchtigte. Alpha-Werte über 0,70 wurden als zufriedenstellend angesehen [25].

Die Prozesse der Anpassung und Validierung des neuen Instruments sind in Abbildung 1 dargestellt.

2.5. Statistische Analyse

Die Normalität der Verteilung der Empathiewerte wurde mit dem Shapiro-Wilk-Test überprüft. Die Sp-SIMCARE-Werte wurden je nach normaler oder schiefer Verteilung der Daten als Mittelwerte und Standardabweichungen oder Median- und Interquartilsbereichswerte zusammengefasst, und kategoriale Variablen wurden als Anteile beschrieben. Der Vergleich der Werte nach Geschlecht, Altersgruppe und simuliertem Szenario erfolgte mit dem Studierenden-t-Test oder dem Mann-Whitney-Test (Vergleich zweier Gruppen) und der ANOVA oder dem Kruskal-Wallis-Test (Vergleich mehrerer Gruppen). Das Signifikanzniveau wurde auf 5% festgelegt ($p < 0,05$). Die Datenanalyse wurde mit dem Statistikpaket Stata v. 15 durchgeführt. Alle identifizierenden Informationen der Studierenden wurden ordnungsgemäß anonymisiert.

3. Ergebnisse

Der kontextuelle Anpassungsprozess stellte sicher, dass die Sp-SIMCARE-Items relevant, klar formuliert, verständlich und für alle Nutzer gleichbedeutend waren. Insgesamt nahmen 95 Studierende, von denen 83% unter 26 Jahre alt und 63,2% weiblich waren, sowie acht SPs an den vier Szenarien zur Validierung des neuen Instruments teil. Aufgrund verschiedener Unwägbarkeiten, die in erster Linie mit der Abwesenheit von Studierenden zusammenhängen, wurden nur 297 der geplanten 380 Begegnungen zwischen Studierenden und SPs durchgeführt (siehe Tabelle 2 mit den demografischen Daten der Studierenden und den Sp-SIMCARE-Werten nach Altersgruppe, Geschlecht und simuliertem Szenario). Von den 297 SP-Bewertungen dieser Begegnungen wiesen 27 (9,1%) drei oder mehr nicht zutreffende oder fehlende Antworten auf und wurden gemäß den von den CARE-Entwicklern empfohlenen Kriterien aus der Analyse ausgeschlossen.

Das Validierungsverfahren umfasste schließlich die Untersuchung der Leistungen der Studierenden in 270 Begegnungen, was zu einer Gesamtpunktzahl von 2.700 Punkten im Sp-SIMCARE führte. Eine Übersicht über die demografischen Daten der Teilnehmer und die Sp-SIMCARE-Werte finden Sie in Tabelle 2. Der Medianwert lag bei 32,5 (IQR, 29,0-37,0), wobei die Werte zwischen 10 und 50 lagen. Die meisten Antworten auf die einzelnen Items der Skala wurden entweder mit „gut“ (45,2%) oder „sehr gut“ (34,1%) bewertet. Es gab keine statistisch signifikanten Unterschiede im Empathieniveau der Teilnehmenden hinsichtlich Alter, Geschlecht oder Art des simulierten Patient*innen (siehe Tabelle 3 mit den Sp-SIMCARE-Werten für die Leistungen der Medizinstudierenden).

Der Sp-SIMCARE-Fragebogen wies eine hohe Augenscheinvalidität auf: Nur 53 (1,96%) der insgesamt 2.700 Ant-

Tabelle 2: Demografische Daten der Medizinstudenten und Sp-SIMCARE-Werte nach Altersgruppe, Geschlecht und simuliertem Szenario

Variable	n (%)	Medianwert (IQR)	p
Altersgruppe			0.57
≤ 25 Jahre	79 (83.2)	32.5 (28.7-36.0)	
> 25 Jahre	16 (16.8)	33.0 (30.5-34.8)	
Geschlecht			0.49
Männlich	35 (36.8)	32.4 (29.5-35.8)	
Weiblich	60 (63.2)	33.5 (28.7-36.4)	
Simulationssituationen			0.08
Akutpatient	74 (27.4)	32.3 (28.0-38.0)	
Chronikpatient	75 (27.8)	30.0 (23.0-39.0)	
Funktionsfähiger Patient	74 (27.4)	34.0 (29.0-37.0)	
Feindseliger Patient *	47 (17.4)	30.0 (23.0-39.0)	

*Signifikante Punktunterschiede ($p \leq 0,001$) zwischen dieser Gruppe und den übrigen simulierten Patientengruppen

worten waren „trifft nicht zu“-Auswahlen oder wurden leer gelassen (siehe Tabelle 2). Von 270 Umfragen wiesen 21 (7,8%) zwei nicht zutreffende oder leere Antworten auf, und 11 (4,1%) wiesen eine solche Antwort auf. Die Begegnungen zeigten, dass solche Antworten am häufigsten bei schwer zu behandelnden Patient*innen auftraten (32/470, 6,81%), während der Anteil bei chronischen Patient*innen deutlich geringer war (19/750, 2,53%) und bei akuten oder funktionellen Patient*innen nicht auftrat (siehe Anhang 1, Tabelle S2). Es wurde festgestellt, dass ein erheblicher Anteil der Befragten entweder „trifft nicht zu“ ausgewählt oder die Antworten auf die Fragen 9 „unterstützt Sie dabei, Verantwortung zu übernehmen“ (26/270, 9,6%) und 10 „arbeitet mit Ihnen zusammen, um einen Aktionsplan zu formulieren“ (23/270, 8,5%) leer gelassen hat. Der Anteil der nicht zutreffenden oder fehlenden Antworten auf die Fragen 7 (4/270, 1,5%) und 8 (3/270, 1,0%) war deutlich geringer. Die Fragen 1 bis 6 wurden in allen Umfragen bewertet.

Es wurde eine signifikante positive Korrelation (Spearman-Rho-Koeffizient, 0,730; $p < 0,001$) zwischen den von den simulierten Patient*innen anhand der Sp-SIMCARE-Skala vergebenen Bewertungen und den Antworten auf die Frage „Ist der Student einfühlsam?“ festgestellt, nachdem die schiefe Verteilung der Bewertungen bestätigt worden war. Die korrigierten Item-Gesamt-Korrelationen lagen zwischen 0,797 und 0,869, und der Cronbach-Alpha-Wert für die Skala betrug 0,960, wobei die Werte für einzelne Items im Bereich von 0,954 bis 0,957 lagen (siehe Tabelle 4 mit Werten für konvergente Validität, Homogenität und interne Reliabilität des Sp-SIMCARE-Fragebogens).

4. Diskussion

Diese Studie liefert Belege für die hohe Validität und Reliabilität eines neuen Fragebogens, der auf dem bekannten CARE-Instrument basiert und die Bewertung der relationalen Empathie im Kontext klinischer Situationen mit

simulierten Patient*innen standardisiert. Die psychometrischen Ergebnisse bestätigen, dass dieses Instrument für den Einsatz bei Medizinstudierenden in solchen Situationen geeignet, valide und potenziell nützlich ist.

Basierend auf den Ergebnissen der Studie zeigte der neue Fragebogen eine hohe Akzeptanz mit nur 1,96% „nicht zutreffenden oder leeren“ Antworten, was mit den 1% vergleichbar ist, die bei der Validierung der vorherigen Sp-CARE-Version beobachtet wurden [24]. Dieses Ergebnis zeigt, dass die in den Simulationsszenarien erzeugten Interaktionen realistisch und vielfältig waren und mit dem neuen Fragebogen effektiv bewertet werden konnten. Frühere Studien, die in anderen Einrichtungen der Primärversorgung durchgeführt wurden, haben ein breites Spektrum an Akzeptanz der ursprünglichen CARE-Messung und abgeleiteten Versionen dokumentiert, wobei einige Arbeiten ähnlich niedrige Raten an nicht zutreffenden oder fehlenden Antworten [16], [24] und andere viel höhere Prozentsätze [14], [17] beschreiben. Die Szenarien mit simulierten Patient*innen, die schwer zu behandeln waren, erhielten den höchsten Prozentsatz an „nicht zutreffenden oder leeren“ Antworten (6,8%). Diese Antworten konzentrierten sich auf zwei bestimmte Punkte: 9 (9,6%) und 10 (8,5%). Diese Prozentsätze stimmten mit ähnlichen Anteilen der Augenscheininvalidität für die Punkte 9 und 10 überein, die in früheren Untersuchungen zur CARE-Messung berichtet wurden [12], [14], [15], [16]. Die geringe Gesichtsgültigkeit der beiden Items war vorhersehbar, da das Management von Emotionen und die Beruhigung von Patient*innen gegenüber der erwarteten Leistung der Studierenden im Umgang mit konfliktreichen Patient*innen überwog. In diesem Zusammenhang wurde in einer früheren Studie sogar postuliert, dass Item 10 bei der Bewertung der relationalen Empathie aus der Bewertung ausgeschlossen werden sollte, da es möglicherweise kein genauer Indikator für die Empathie eines Arztes/einer Ärztin ist, sondern eher die gemeinsame Entscheidungsfindung widerspiegelt [26].

Der Medianwert des Sp-SIMCARE-Scores unserer Studienpopulation (32,5) lag deutlich unter den zuvor in CARE-

Tabelle 3: Sp-SIMCARE-Werte für die Leistung von Medizinstudenten

Item	Schlecht n (%)	Befriedigend n (%)	Gut n (%)	Sehr gut n (%)	Ausgezeichnet n (%)	Gesamt n (%)	n (%)
1. Ihnen ein Gefühl der Geborgenheit vermitteln	9 (3.3)	25 (9.3)	118 (43.7)	104 (38.5)	14 (5.2)	0 (0.0)	270 (100.0)
2. Sie Ihre „Geschichte“ erzählen lassen	1 (0.4)	20 (7.4)	120 (44.4)	117 (43.3)	12 (4.4)	0 (0.0)	270 (100.0)
3. Ihnen wirklich zuhören	4 (1.5)	22 (8.1)	108 (40.0)	125 (46.3)	11 (4.1)	0 (0.0)	270 (100.0)
4. Sich für Sie als ganzen Menschen interessieren	5 (1.8)	35 (13.0)	119 (44.1)	102 (37.8)	9 (3.3)	0 (0.0)	270 (100.0)
5. Ihre Sorgen vollständig verstehen	8 (3.0)	45 (16.7)	133 (49.3)	75 (27.8)	9 (3.3)	0 (0.0)	270 (100.0)
6. Fürsorge und Mitgefühl zeigen	7 (2.6)	38 (14.1)	130 (48.1)	87 (32.2)	8 (3.0)	0 (0.0)	270 (100.0)
7. Positiv sein	7 (2.6)	26 (9.6)	124 (45.9)	100 (37.0)	9 (3.3)	4 (1.5)	270 (100.0)
8. Dinge klar erklären	5 (1.8)	40 (14.8)	128 (47.4)	87 (32.2)	10 (3.7)	3 (1.0)	270 (100.0)
9. Ihnen helfen, die Kontrolle zu übernehmen	7 (2.6)	49 (18.1)	127 (47.0)	54 (20.0)	7 (2.6)	26 (9.6)	270 (100.0)
10. Gemeinsam mit Ihnen einen Aktionsplan erstellen	9 (3.3)	46 (17.0)	114 (42.2)	69 (25.6)	9 (3.3)	23 (8.5)	270 (100.0)

Tabelle 4: Konvergente Validität, Homogenität und interne Reliabilität des Sp-SIMCARE-Fragebogens

Item	Konvergente Validität*	Skalenmittelwert bei Löschung des Items	Korrigierte Item-Gesamt-Korrelation	Cronbachs Alpha
1. Ihnen ein Gefühl der Geborgenheit vermitteln	0.735	29.00	0.884	0.955
2. Sie Ihre „Geschichte“ erzählen lassen	0.590	28.89	0.827	0.957
3. Ihnen wirklich zuhören	0.590	28.90	0.871	0.955
4. Sich für Sie als ganzen Menschen interessieren	0.589	29.06	0.870	0.955
5. Ihre Sorgen vollständig verstehen	0.607	29.21	0.848	0.957
6. Fürsorge und Mitgefühl zeigen	0.685	29.14	0.896	0.954
7. Positiv sein	0.644	29.05	0.850	0.956
8. Dinge klar erklären	0.611	29.12	0.856	0.956
9. Ihnen helfen, die Kontrolle zu übernehmen	0.593	29.34	0.847	0.957
10. Gemeinsam mit Ihnen einen Aktionsplan erstellen	0.553	29.28	0.840	0.957

* Alle Korrelationen sind statistisch signifikant bei einem p-Wert <0,001

Validierungs- oder Implementierungsstudien in europäischen Primärversorgungseinrichtungen veröffentlichten Mittel- oder Medianwerten (über 40) [12], [16]. Die bemerkenswerten Unterschiede in den Werten zwischen unserer und anderen europäischen Studien lassen sich durch die unterschiedlichen untersuchten Populationen erklären: Unsere Studie bewertete die Leistung von Medizinstudierenden im Grundstudium, die nur begrenzte Erfahrungen im Umgang mit simulierten Patient*innen hatten, während andere Studien die Leistung von Hausärzt*innen bewerteten, die regelmäßig und intensiv mit ihren Patient*innen interagieren und daher eher empathische Kompetenzen zeigen. Interessanterweise wurden keine signifikanten Unterschiede bei den Sp-SIMCARE-Werten in Abhängigkeit vom Geschlecht der Medizinstudierenden festgestellt. Dieses Ergebnis steht im Einklang mit früheren Studien, die darauf hindeuten, dass die Werte der CARE-Messung weder durch das Geschlecht der medizinischen Fachkräfte noch durch die Merkmale der Konsultation wesentlich beeinflusst wurden [13], [14], [15], [24].

Die vorliegende Arbeit weist logischerweise einige methodische Stärken und Einschränkungen auf. Zu den Stärken zählt insbesondere der sorgfältige sequenzielle Prozess, der zur Anpassung von Sp-CARE an den Simulationskontext durchgeführt wurde. Darüber hinaus lässt die Tatsache, dass das Tool in verschiedenen pathologischen Szenarien und mit simulierten Patient*innen mit unterschiedlichen Profilen getestet wurde, auf eine gute Benutzer*innenfreundlichkeit schließen [27]. In diesem Sinne berücksichtigten die Szenarien, die zuvor einem Entwurfs- und Validierungsprozess durch einen Expertenausschuss unterzogen worden waren, nicht nur verschiedene klinische Situationen, sondern auch verschiedene Patient*inentypen, darunter ein spezifisches Szenario mit feindseligen SPs, die sich durch mangelnde Kooperationsbereitschaft oder hohe Aggressivität auszeichneten. Eine Einschränkung der Studie ist das Fehlen eines objektiven

Goldstandards für relationale Empathie. Um die konvergente Validität des Sp-SIMCARE-Fragebogens zu bewerten, war es daher notwendig, seine Ergebnisse mit einem Proxy für den Goldstandard zu vergleichen. In diesem Fall haben wir keine validierte Umfrage zur Messung von Empathie verwendet, sondern uns stattdessen auf die selbst berichtete Wahrnehmung von Empathie durch den simulierten Patient*innen während der klinischen Begegnung gestützt. Dies diente als unser externer Standard oder Kriterium. Auch wenn dies von manchen als Einschränkung angesehen werden mag, handelt es sich doch um ein in ähnlichen Fällen häufig verwendetes Verfahren. Andererseits war die Stichprobengröße der Studie klein, was die Möglichkeit explorativer und konfirmatorischer Faktorenanalysen einschränkte.

Darüber hinaus legt die Herkunft der Studienpopulation, die sich aus freiwilligen Studierenden im vierten Studienjahr zusammensetzt, nahe, dass bei der Verallgemeinerung der Ergebnisse Vorsicht geboten ist.

Der Fokus des Sp-SIMCARE-Fragebogens auf die relationalen Aspekte von Empathie bietet einen Vorteil. Er bewertet, wie sich Empathie in simulierten Interaktionen in konkrete Handlungen umsetzt, und liefert damit wertvolle Erkenntnisse, die über ein subjektives und emotionales Verständnis hinausgehen. Der Fragebogen ist besonders sensibel für die Komplexität der Interaktion mit unterschiedlichen Szenarien und Patient*innen, einschließlich der schwierigsten Fälle. Anpassungsfähigkeit ist in der medizinischen Ausbildung von entscheidender Bedeutung, insbesondere wenn es darum geht, Empathie in schwierigen klinischen Situationen, z. B. mit konfliktreichen Patient*innen, zu testen. Hervorzuheben ist, dass die Verwendung simulierter Umgebungen für die Validierung des Fragebogens eine strategische Entscheidung ist. Sie bietet eine kontrollierte und sichere Umgebung und gewährleistet eine konsistente und faire Bewertung für alle Studierenden, insbesondere für diejenigen, die sich noch in der Ausbildung befinden. Diese Methodik

beseitigt die praktischen Hindernisse, die mit dem Einholen von direktem Feedback von Patient*innen verbunden sind, indem sie eine strukturierte und reproduzierbare Bewertung ermöglicht. Wichtig ist, dass die Validierung von Sp-SIMCARE bei fortgeschrittenen Medizinstudierenden auf seine potenzielle Nützlichkeit für die Bewertung der relationalen Empathie in anderen Berufsgruppen des Gesundheitswesens hindeutet, obwohl weitere Studien erforderlich sind, um dies zu bestätigen.

5. Fazit

Zusammenfassend lässt sich sagen, dass sich der Sp-SIMCARE-Fragebogen als psychometrisch valide und zuverlässige Methode zur Bewertung von Medizinstudierenden durch simulierte Patient*innen erwiesen hat. Die Einzigartigkeit des Fragebogens liegt in seiner Fähigkeit, die relationale Dimension der Empathie zu messen, wodurch er ein praktisches Instrument zur Bewertung dieser Kompetenz darstellt. Der Einsatz dieses neuen Instruments könnte potenziell bei der Konzeption und Umsetzung von Maßnahmen zur Förderung der Empathie bei angehenden Ärzt*innen während ihrer gesamten Ausbildung hilfreich sein.

Danksagung

Die Autor*innen danken den simulierten Patient*innen und den Studierenden des Studiengangs Medizin der Universität International de Catalunya, die an dieser Studie teilgenommen haben, für ihren Beitrag.

Anmerkungen

Ethische Genehmigung und Einwilligung zur Teilnahme

Das Studienprotokoll wurde von der Ethikkommission des Studienzentrums genehmigt (Registrierung MED-2022-07). Durch die Genehmigung des Protokolls entfiel die Verpflichtung, eine Einverständniserklärung von den Studienteilnehmenden einzuholen.

Finanzierung

Diese Studie wurde teilweise durch einen wettbewerbsorientierten Zuschuss für medizinische Bildungsforschungsprojekte finanziert, der von der Sociedad Española de Educación Médica (SEDEM) vergeben wurde.

Autorenschaft

Pedro Brotos de los Reyes und Montserrat Virumbrales Cancio teilen sich die Erstautorenschaft.

Beiträge der Autor*innen

- PB: Datenanalyse, Interpretation der Ergebnisse, Erstellung des Manuskriptentwurfs, endgültige Freigabe des Manuskripts.
- MV: Studienkonzeption und -design, Datenerhebung, Interpretation der Ergebnisse, endgültige Freigabe des Manuskripts.
- PC: Interpretation der Ergebnisse, endgültige Freigabe des Manuskripts
- XM: Datenerhebung, Interpretation der Ergebnisse, endgültige Freigabe des Manuskripts
- SD: Datenerhebung, Interpretation der Ergebnisse, endgültige Freigabe des Manuskripts
- AB: Studienkonzeption und -design, Studienmanagement, Interpretation der Ergebnisse, endgültige Freigabe des Manuskripts

ORCIDs der Autor*innen

- Pedro Brotos de los Reyes: [0000-0002-2399-7320]
- Montserrat Virumbrales Cancio: [0000-0003-3541-2948]
- Pere Castellvi: [0000-0002-3920-8576]
- Xavier Martinez-Regada: [0009-0009-5013-7738]
- Albert Balaguer: [0000-0002-5222-8635]

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Anhänge

Verfügbar unter <https://doi.org/10.3205/zma001844>

1. Anhang_1.pdf (321 KB)
Ergänzendes Material

Literatur

1. Derksen F, Olde Hartman TC, van Dijk A, Plouvier A, Bensing J, Lagro-Janssen A. Consequences of the presence and absence of empathy during consultations in primary care: a focus group study with patients. *Patient Educ Couns.* 2017;100(5):987-993. DOI: 10.1016/j.pec.2016.12.003
2. Hojat M, LaNoue M. Exploration and confirmation of the latent variable structure of the Jefferson scale of empathy. *Int J Med Educ.* 2014;20(5):73-81. DOI: 10.5116/ijme.533f.0c41
3. Mercer SW, Reynolds WJ. Empathy and quality of care. *Br J Gen Pract.* 2002;52 Suppl(Suppl):S9-S12.
4. Hall JA, Schwartz R, Duong F, Niu Y, Dubey M, DeSteno D, Sanders JJ. What is clinical empathy? Perspectives of community members, university students, cancer patients, and physicians. *Patient Educ Couns.* 2021;104(5):1237-1245. DOI: 10.1016/j.pec.2020.11.001
5. Schwartz R, Dubey M, Blanch-Hartigan D, Sanders JJ, Hall JA. Physician empathy according to physicians: A multi-specialty qualitative analysis. *Patient Educ Couns.* 2021;104(10):2425-2431. DOI: 10.1016/j.pec.2021.07.024

6. Sanders JJ, Dubey M, Hall JA, Catzen HZ, Blanch-Hartigan D, Schwartz R. What is empathy? Oncology patient perspectives on empathic clinician behaviors. *Cancer*. 2021;127(22):4258-4265. DOI: 10.1002/cncr.33834
7. Decety J, Svetlova M. Putting together phylogenetic and ontogenetic perspectives on empathy. *Dev Cogn Neurosci*. 2012;2(1):1-24. DOI: 10.1016/j.dcn.2011.05.003
8. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Heal Prof*. 2022;27(3):237-251. DOI: 10.1177/0163278704267037
9. Del Canale S, Louis DZ, Maio V, Wang X, Rossi G, Hojat M, Gonnella JS. The relationship between physician empathy and disease complications: an empirical study of primary care physicians and their diabetic patients in Parma, Italy. *Acad Med*. 2012;87(9):1243-1249. DOI: 10.1097/ACM.0b013e3182628fbf
10. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' empathy and clinical outcomes for diabetic patients. *Acad Med*. 2011;86(3):359-364. DOI: 10.1097/ACM.0b013e3182086fe1
11. Hemmerdinger JM, Stoddart SD, Lilford RJ. A systematic review of tests of empathy in medicine. *BMC Med Educ*. 2007;7:24. DOI: 10.1186/1472-6920-7-24
12. Mercer SW, Maxwell M, Heaney D, Watt GC. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract*. 2004;21(6):699-705.
13. Mercer SW, McConnachie A, Maxwell M, Heaney D, Watt GC. Relevance and practical use of the Consultation and Relational Empathy (CARE) Measure in general practice. *Fam Pract*. 2005;22(3):328-334. DOI: 10.1093/fampra/cmh621
14. Fung CS, Hua A, Tam L, Mercer SW. Reliability and validity of the Chinese version of the CARE measure in a primary care setting in Hong Kong. *Fam Pract*. 2009;26(5):398-406. DOI: 10.1093/fampra/cmp044
15. Aomatsu M, Abe H, Abe K, Yasui H, Suzuki T, Sato J, Ban N, Mercer SW. Validity and reliability of the Japanese version of the CARE measure in a general medicine outpatient setting. *Fam Pract*. 2014;31(1):118-126. DOI: 10.1093/fampra/cmt053
16. van Dijk I, Scholten Meilink Lenferink N, Lucassen PL, Mercer SW, van Weel C, Hartman TC, Speckens AE. Reliability and validity of the Dutch version of the Consultation and Relational Empathy Measure in primary care. *Fam Pract*. 2017;34(1):119-124. DOI: 10.1093/fampra/cmw116
17. Hanževački M, Jakovina T, Bajić Ž, Tomac A, Mercer S. Reliability and validity of the Croatian version of Consultation and Relational Empathy (CARE) Measure in primary care setting. *Croat Med J*. 2015;56(1):50-56. DOI: 10.3325/cmj.2015.56.50
18. Solaja O, Moloo H, Hopkins E, Khan N, Gilbert RW, Neville A, Kubelik D, Maziak D, Rowe N, Odell M, Momtazi M. Implementation, results and face validity of the Consultation and Relational Empathy measure in a Canadian department of surgery. *Can J Surg*. 2022;65(6):E749-E755. DOI: 10.1503/cjs.003721
19. Del Barrio LG, Rodríguez-Díez C, Gea A, Arbea L, Pereira J, Díez N. Impact of a longitudinal course on medical professionalism on the empathy of medical students. *Patient Educ Couns*. 2024;119:108042. DOI: 10.1016/j.pec.2023.108042
20. Duran A, Donelan C, Bowman Peterson J, Gladding SP, Weissmann P, Roth CS. Communicating value to patients—a high-value care communication skills curriculum. *Postgrad Med*. 2021;133(2):231-236. DOI: 10.1080/00325481.2020.1807728
21. Schreckenbach T, Ochsendorf F, Sterz J, Rüsseler M, Bechstein WO, Bender B, Bechtoldt MN. Emotion recognition and extraversion of medical students interact to predict their empathic communication perceived by simulated patients. *BMC Med Educ*. 2018;18(1):237. DOI: 10.1186/s12909-018-1342-8
22. Vogel D, Meyer M, Harendza S. Verbal and non-verbal communication skills including empathy during history taking of undergraduate medical students. *BMC Med Educ*. 2018;18(1):157. DOI: 10.1186/s12909-018-1260-9
23. Buffel du Vaure C, Lemogne C, Bunge L, Catu-Pinault A, Hoertel N, Ghasarossian C, Vincens ME, Galam E, Jaury P. Promoting empathy among medical students: A two-site randomized controlled study. *J Psychosom Res*. 2017;103:102-107. DOI: 10.1016/j.jpsychores.2017.10.008
24. García Del Barrio L, Rodríguez-Díez C, Martín-Lanas R, Costa P, Costa MJ, Díez N. Reliability and validity of the Spanish (Spain) version of the consultation and relational empathy measure in primary care. *Fam Pract*. 2021;38(3):353-359. DOI: 10.1093/fampra/cmaa135
25. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. New York (NY): Oxford University Press; 2008.
26. Wirtz M, Boecker M, Forkmann T, Neumann M. Evaluation of the "Consultation and Relational Empathy" (CARE) measure by means of Rasch-analysis at the example of cancer patients. *Patient Educ Couns*. 2011;82(3):298-306. DOI: 10.1016/j.pec.2010.12.009
27. Brotos P, Virumbrales M, Elorduy M, Díaz de Castellví S, Mezquita P, Gené E, Balaguer A. Improvement of medical students' performance in simulated patient interviews by pre-clinical communication training. *Int J Med Educ*. 2022;13:148-153. DOI: 10.5116/ijme.6299.c15f

Korrespondenzadresse:

Albert Balaguer
 Universitat Internacional de Catalunya, School of Medicine and Health Sciences, Department of Medicine, Josep Trueta Street, s/n. ES-08195 Sant Cugat del Vallès, Spanien
 abalaguer@uic.es

Bitte zitieren als

Brotos de los Reyes P, Virumbrales Cancio M, Castellvi P, Martínez-Regada X, Balaguer A. Assessment of empathy by simulated patients: Adaptation and validation of a new instrument. *GMS J Med Educ*. 2026;43(4):Doc50. DOI: 10.3205/zma001844, URN: urn:nbn:de:0183-zma0018446

Artikel online frei zugänglich unter

<https://doi.org/10.3205/zma001844>

Eingereicht: 03.10.2024

Überarbeitet: 08.06.2025

Angenommen: 27.10.2025

Veröffentlicht: 15.04.2026

Copyright

©2026 Brotos de los Reyes et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.