

# Simulated patients' role-portrayal in the clinical skills part of the Swiss federal licensing exam is of high quality and improves further over time as measured with the FAIR OSCE instrument

## Abstract

**Introduction:** Simulation-based teaching and assessment are integral to education in the health professions, with simulated patients (SP) being a widely accepted strategy. Ensuring high-quality SP role-portrayal is crucial for the authenticity and standardization of assessments, particularly in high-stakes exams like the Swiss Federal Licensing Examination (FLE).

The study assesses the quality of SP role-portrayal over consecutive instances of the Swiss FLE. We hypothesized that the quality of role-portrayal improves over time.

**Methods:** The study employed the FAIR OSCE instrument to assess SP role-portrayal in five consecutive FLE exams from 2016 to 2021. The instrument, developed between 2011 and 2014, includes four categories: introduction, delivery of information, portrayal, and others. Data analysis involved retrospective examination of FAIR OSCE ratings, calculating item scores, and overall mean scores for each exam year.

**Results:** The study involved 37 SP educators observing 1803 SP-candidate interactions across five exam sites. Results demonstrated a continuous improvement in SP role-portrayal over the five-year period, with significant differences between 2016 and subsequent years. The overall mean scores of SP role-play ratings increased steadily, indicating a positive trend in SP performance.

**Discussion:** The findings supported the hypothesis of continuous improvement of SP role-portrayal within the Swiss FLE. The quality of role portrayal not only improved consistently but also maintained a very high level, with no items on the FAIR OSCE instrument rated "do not agree" more than 5% of the time. This suggests that SPs role play aligned well with case scripts, reflecting the intended authenticity and standardization of assessments.

Limitations were acknowledged, including potential bias in local SP educators rating their own SPs and the study's sole focus on SP role-portrayal. While the findings contribute to understanding SP effectiveness in standardized, high-stakes clinical exams, the study did not scrutinize other potential sources of variance.

**Conclusion:** In conclusion, the research demonstrated a continuous improvement and high quality of SP role-portrayal in the Swiss FLE over five years. Well-trained SPs, assessed using the FAIR OSCE instrument, play a crucial role in maintaining the standardized and high-quality nature of clinical skills exams in a high-stakes context. Further research could explore additional factors influencing overall exam quality and address potential biases in SP educator ratings.

**Keywords:** OSCE, objective structured clinical examination, standardization, standardized patient, simulated patient, high-stakes examinations, licensing exam, patient simulation

Kai P. Schnabel<sup>1</sup>

Daniel Bauer<sup>1</sup>

Felix M. Schmitz<sup>1</sup>

Tanja Hitzblech<sup>1</sup>

Beate G. Brem<sup>1</sup>

<sup>1</sup> University of Bern, Institute for Medical Education, Bern, Switzerland

## Introduction

Simulation-based teaching and assessment are crucial in health professions education and assessment [1], [2]. Among different simulation-related options such as the use of mannequins, full-scale simulators and computer-based applications, employing simulated patients (aka simulated participants, SP) is a proven strategy for delivering simulation scenarios [2], [3]. First introduced in the 1960s [4], [5], working with SPs in learning and assessment contexts is nowadays widely accepted and applied by many medical schools all over the world [6]. Objective Structured Clinical Examinations (OSCEs) [7] regularly employ SPs, in which medical students interact with SPs portraying pre-defined patient cases, while examiners rate the students' performance. The quality of the SPs' role-portrayal is of critical importance to ensure fair and valid exams, above all in terms of authenticity and standardization [8]: while an *insufficiently authentic* role-portrayal limits the representativeness of the task put before the student, *insufficiently standardized* role-portrayal limits the comparability of their tasks and thus exam fairness, both are threats to the exam's validity [2], [9], [10]. It is the role of SP educators to ensure their role-portrayal has a high level and good balance of authenticity and standardization [11], [12].

Switzerland introduced the new Federal Licensing Examination in (FLE) in 2011 [13], with a written exam and an OSCE that graduates of the medical Masters programs take in order to achieve licensing as physicians. An increasing number of candidates (medical students and candidates from non-European countries, seeking licensure in Switzerland) have to be assessed in the FLE with over a thousand taking the test in 2021 [14]. Several measures were put in place to ensure the highest quality in the development and delivery of this assessment [15], [16], [17]. The quality of the overall FLE is mirrored, e.g., with Cronbach's alpha of above 0.84 as indicator of measurement reliability) [17].

To ensure each and every scenario's SPs with their diverse individual backgrounds and acting experience are amalgamated into homogenous groups that portray their roles well over the whole exam day, a protocol for training and quality assurance was put in place. This includes 5 days of standardization meetings on a federal level to establish a common understanding of the scripts and scenarios and how to put them into practice and generating videographic material depicting the prototypical realization of every scenario (e.g., standardization of pain reactions, application of moulage [18], [19]). SP educators and SPs from their local program then use that material in two spaced trainings session of 2 hrs each to ensure authentic and standardized role portrayal during the exam [17]. During the exam, quality of SPs' role portrayal (and other aspects of the exam) is controlled via live observation, e.g., via one-way mirrors or via audio-visual feed into a control room.

Relying on an assessment format as complex and resource-intensive as the OSCE to make decisions that can

impact both patients' safety (candidates that pass as false-positive) but also test-takers' careers (candidates that fail as false-negative) warrants an analysis to see if SP role-portrayal in high-stakes OSCE can be performed at high levels and continuously. Thus, the aim of the present study was to assess the quality of simulated patients' (SP) role-portrayal in Swiss high-stakes clinical exams over time. We hypothesized that the quality of SP role-portrayal in the Swiss FLE continues to improve over time.

## Methods

In order to determine if SP role-portrayal in high-stakes exams is performed at high levels in general and over time, we conducted a cohort study with five Swiss SP programs involved in the delivery of the OSCE within the FLE.

## Setting

The OSCEs were conducted at the time of this analysis at 5 sites (Basel, Bern, Geneva, Lausanne and Zurich), simultaneously in French and German, over three days, each run covering 12 stations, and almost all of which employed SPs (comprehensive data about the exam is available in attachment 1).

## Procedure

To determine if OSCE-related SP role-portrayal is performed at high levels and consistently so, SP role-portrayal was assessed during the OSCEs within the FLE from 2016 to 2021, representing five consecutive exams as the 2020 clinical skills exam was not administered due to the COVID-19 pandemic. These five exams assessed over 5000 candidates' clinical competencies, for which SPs were employed to deliver 70 parallel runs in total at 12 stations each. Data on these SPs' role-portrayal were collected at all five sites over the five exams in random observations, via direct observation or live-stream as measured with the FAIR OSCE instrument. Raters were the local SP trainer (1-5 at each site) and one federal observer. The SP trainers were either physicians or experienced health professionals or artists with 2-20 years' experience in the field and members of the federal SP trainer team.

## Instrument

The FAIR OSCE instrument was developed between 2011 and 2014 by an interfaculty board of Swiss SP educators attempting to assess SP role-portrayal using *case-independent* observable criteria. It is based on the available literature and underwent a consensus-building process among the interfaculty board, supporting its content validity [15]. The instrument was additionally used in SP trainings as well as delivery of local formative and sum-

mative OSCEs at the faculty of Bern and has proven its feasibility and usability [15], [20].

The instrument encompasses four categories:

1. introduction (3 items),
2. delivery of information (7 items),
3. portrayal (7 items), and
4. others (3 items).

The items are assessed on a scale with the points  $-1=do\ not\ agree$ ,  $+1=agree$ ,  $+2=completely\ agree$ , and  $not\ applicable$ . (cf. attachment 2)

## Data analysis

FAIR OSCE ratings from 2016 to 2021 were analysed retrospectively. Item scores were calculated to highlight the quality of SP role-portrayal within each year's exam, and overall mean scores were calculated for each of the 5 exams (2016 to 2021, items rated as "not applicable" were excluded), i.e. all ratings derive from all sites administering the examination in a given year.

Statistical differences between examination years were calculated with one-way ANOVA. Effect sizes were calculated using eta squared with  $\eta^2 \sim 0.01$  for small,  $\eta^2 \sim 0.06$  for medium and  $\eta^2 \geq 0.14$  for large effects. Post hoc tests were calculated using Bonferroni tests for each possible group comparison.

## Ethics approval

Ethics approval for this study was not necessary in accordance with Swiss legislation (chapter 1, article 2 of the 2011 Swiss Human Research Act). Ethics principles according to the WMA Helsinki declaration and good scientific practices were followed throughout.

## Results

During the 5 OSCEs between 2016 and 2021, totalling  $N=37$  different SP educators (approximately 2-4 per site and year, mostly the same SP educators over the years) observed a total of  $N=1803$  SP-candidate interactions (Basel  $n=247$ , Bern  $n=419$ , Geneva  $n=357$ , Lausanne  $n=467$ , and Zurich  $n=313$ ) and rated the SP role-portrayal.

The number of observations increased from an initial 168 observations in 2016 to a maximum of 499 observations in 2021.

To exemplify, table 1 shows the rating results for the OSCE 2021, based on 499 independent observations.

The overall mean scores of the SP role-play ratings increased continuously over the years (see table 2). The ANOVA showed that at least one increase across the years was of statistical significance ( $F=9.2$ ,  $p<0.001$ ;  $\eta^2=0.02$ ). The appropriate Bonferroni post hoc tests revealed that the results from 2018-2021 were significantly better compared to 2016 (cf. table 2).

## Discussion

In terms of our hypothesis, we were able to demonstrate that the quality of the SP role-portrayal increased continuously over the 5 years analyzed, significantly so when comparing data from 2021 versus 2016 and 2017, 2018 and 2017 versus 2016.

This means we were able to confirm the initial hypothesis of increasing quality.

The quality of the role portrayal not only improved continuously but also was at a very high level with no items on the FAIR OSCE instrument rated "do not agree" more often than in 5% of random observations in all four categories, implying role-portrayal was as intended by the case scripts. Both these results can be seen as indicators both the exam and SPs employed are of the highest quality. It is also an indicator that the FAIR OSCE instrument is able to measure the increasing quality of the exam even though the measurement already started at a very high level considering the starting point of the FLE in 2011 (when one would expect the steepest learning curve), adding to the instrument's validity argument.

The findings are incidentally not a measure of the overall quality of the FLE but a mosaic, as other potential sources of variance like the examiners, checklists and further contextual factors were not scrutinized and remained beyond the scope of this study. A potential source of bias in the study at hand lies in how local SP educators rated the quality of their own SPs' role-portrayal, who were trained by themselves or their colleagues. This however constitutes a systematic error that would apply to all sites. Inter-rater agreement using the FAIR OSCE has so far not been reported and was beyond the scope of this study but should be addressed in future studies. Parallel assessment with the authenticity items of the MaSP [21] could contribute more to the validity argument of the FAIR OSCE. While the study was conducted in only one country in the context of only one exam, the fact it was performed in the context of a highly-standardized high-stakes federal licensing exam performed at five sites in two languages gives it a certain merit.

## Conclusions

SPs' role-portrayal in a series of highly standardized high-stakes exams was shown to be of highest quality and increasing consecutively over all exam cohorts analyzed. Well-trained SPs observed with a valid instrument thus contribute to a standardized, high-quality clinical skills exam.

**Table 1: FAIR OSCE response rates as derived from the 2021 exam**

Item	Response rates [%]			
	Agree completely	Agree	Do not agree	Not applicable
<i>Introduction</i>				
<i>First sentence verbatim</i>	92.2%	6.4%	0.6%	0.8%
<i>Second sentence only after pause/prompt</i>	49.3%	13.6%	3.8%	33.3%
<i>Second sentence verbatim</i>	57.1%	8.2%	2.6%	32.1%
<i>Standardization</i>				
<i>Correct facts</i>	86.2%	13.2%	0.4%	0.2%
<i>Correct portion of Information</i>	83.8%	14.8%	1.2%	0.2%
<i>Correct answers to open questions</i>	76.2%	13.0%	0.6%	10.2%
<i>Correct delivery of prompts</i>	86.0%	12.0%	0.6%	1.4%
<i>Correct improvisation</i>	67.5%	8.2%	0.2%	24.0%
<i>No spontaneous information</i>	86.6%	11.2%	1.8%	0.4%
<i>No spontaneous questions</i>	93.0%	4.8%	2.0%	0.2%
<i>Role-Portrayal</i>				
<i>Correct verbal expression</i>	91.2%	8.4%	0.4%	0.0%
<i>Correct non-verbal expression</i>	90.0%	9.8%	0.2%	0.0%
<i>No cues</i>	91.8%	7.8%	0.0%	0.4%
<i>Correct portrayal of disease symptoms</i>	87.0%	12.0%	0.2%	0.8%
<i>Correct emotions</i>	86.2%	11.8%	0.8%	1.2%
<i>Like a real patient</i>	91.6%	8.2%	0.0%	0.2%
<i>Stays in role until candidate leaves</i>	87.8%	9.2%	0.8%	2.2%
<i>Others</i>				
<i>Correct clothing</i>	94.4%	2.4%	0.8%	2.4%
<i>Correct moulage</i>	23.4%	0.2%	0.0%	76.4%
<i>Handling material with care</i>	27.1%	1.0%	0.0%	71.9%

Note: item titles are abbreviated

**Table 2: Statistics related to the FAIR OSCE scores per examination year**

Year	Descriptives of the FAIR OSCE scores			
	Number	Mean	Standard deviation	Standard error of measurement
2016*	168	1.748	0.36	0.02
2017**	267	1.811	0.27	0.02
2018	407	1.851	0.26	0.01
2019	462	1.853	0.23	0.01
2021	499	1.869	0.19	0.01

Note: \* Bonferroni post hoc tests revealed statistically significant differences between the scores from examination year 2016 vs. 2018, 2019, and 2021 (all accordant  $p < 0.001$ ); \*\* Bonferroni post-hoc tests revealed a statistically significant difference between the scores from examination year 2017 vs. 2021 ( $p < 0.05$ ).

## Authors' ORCIDs

- Kai P. Schnabel: [0000-0002-6977-2717]
- Daniel Bauer: [0000-0002-3337-3327]
- Tanja Hitzblech: [0000-0003-0876-0373]
- Beate G. Brem: [0000-0002-0551-9587]

## Acknowledgements

The authors would like to extend their gratitude towards all SP educators applying the instrument and collating the data during the exams that served as basis for the analyses.

## Competing interests

The authors declare that they have no competing interests.

## Attachments

Available from <https://doi.org/10.3205/zma001736>

1. Attachment\_1.pdf (139 KB)  
Comprehensive description of the clinical skills part of the Swiss Federal Licensing Exam in Medicine
2. Attachment\_2.pdf (247 KB)  
English version of the FAIR OSCE instrument in its current form

## References

1. Bradley P. The history of simulation in medical education and possible future directions. *Med Educ.* 2006;40(3):254-62. DOI: 10.1111/j.1365-2929.2006.02394.x
2. Bauer D, Lahner FM, Huwendiek S, Schmitz FM, Guttormsen S. An overview of and approach to selecting appropriate patient representations in teaching and summative assessment in medical education. *Swiss Med Wkly.* 2020;150:w20382. DOI: 10.4414/smw.2020.20382
3. Nestel D, Kneebone R. Perspective: authentic patient perspectives in simulations for procedural and surgical skills. *Acad Med.* 2010;85(5):889-893. DOI: 10.1097/ACM.0b013e3181d749ac
4. Barrows HS, Abrahamson S. The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. *J Med Educ.* 1964;39:802-5.
5. Barrows HS. Simulated patients in medical teaching. *Can Med Assoc J.* 1968;98(14):674-676.
6. Nestel D, Bearman M. *Simulated Patient Methodology.* Malden (MA); Oxford (UK): West Sussex UK: John Wiley & Sons Ltd.; 2015. DOI: 10.1002/9781118760673
7. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
8. Smith CM, Gliva-Mcconvey G, Chapin A. Training SPs for Authentic Role Portrayal. In: Gliva-McConvey G, Nicholas CF, Clark L, editors. *Comprehensive Healthcare Simulation: Implementing Best Practices in Standardized Patient Methodology Comprehensive Healthcare Simulation.* Cham: Springer; 2020. DOI: 10.1007/978-3-030-43826-5
9. Haladyna TM, Downing SM. Construct-Irrelevant Variance in High-Stakes Testing. *Educ Measurement.* 2004;23(1):17-27. DOI: 10.1111/j.1745-3992.2004.tb00149.x
10. Smith CM, O'Byrne CO, Nestel D. Simulated patient methodology and assessment. In: Nestel D, Bearman M, editors. *Simulated Patient Methodology.* Malden (MA); Oxford (UK): West Sussex UK: John Wiley & Sons Ltd.; 2015. p.85-92. DOI: 10.1002/9781118760673.ch12
11. Lewis KL, Bohnert CA, Gammon WL, Hölzer H, Lyman L, Smith C, Thompson TM, Wallace A, Gliva-Mcconvey G. The Association of Standardized Patient Educators (ASPE) Standards of Best Practice (SOBP). *Adv Simul (Lond).* 2017;2:10. DOI: 10.1186/s41077-017-0043-4
12. Brem B, Steinweg B. Prüfungen mit SPs. In: Peters T, Thrien C, editor. *Simulationspatienten Handbuch für die Aus- und Weiterbildung in medizinischen und Gesundheitsberufen.* 1. Auflage. Göttingen: Hogrefe Verlag; 2018.
13. Guttormsen S, Beyeler C, Bonvin R, Feller S, Schirlo C, Schnabel K, Schurter T, Berendonk C. The new licencing examination for human medicine: from concept to implementation. *Swiss Med Wkly.* 2013;143:w13897. DOI: 10.4414/smw.2013.13897
14. Bundesamt für Gesundheit. Statistiken Ärztinnen/Ärzte. Bern: Bundesamt für Gesundheit; 2022. Zugänglich unter/available from: [https://www.bag.admin.ch/dam/bag/de/dokumente/berufe-gesundheitswesen/medizinalberufe/statistiken/med/aerztinnen-aerzte-2021.pdf.download.pdf/Aerztinnen\\_und\\_Aerzte\\_2021\\_DE.pdf](https://www.bag.admin.ch/dam/bag/de/dokumente/berufe-gesundheitswesen/medizinalberufe/statistiken/med/aerztinnen-aerzte-2021.pdf.download.pdf/Aerztinnen_und_Aerzte_2021_DE.pdf)
15. Brem B, Richter S, Schnabel KP. FAIR\_OSCE - Feedbackstructure for assessment of interactive roleplay in Objective Structured Clinical Exams. In: 4th Swiss Conference on Standardized Patients and Simulation in Health Care. Bern, 10-12 September 2014. SC2-5. Zugänglich unter/available from: <https://with-simulation.ch/programm-2014/>
16. Bachmann C, Kropf R, Biller S, Schnabel KP, Junod Perron N, Monti M, Berendonk C, Huwendiek S, Breckwoldt J. Development and national consensus finding on patient-centred high stakes communication skills assessments for the Swiss Federal Licensing Examination in Medicine. *Patient Educ Couns.* 2021;104(7):1765-1772. DOI: 10.1016/j.pec.2020.12.003
17. Berendonk C, Schirlo C, Balestra G, Bonvin R, Feller S, Huber P, Junger E, Monti M, Schnabel K, Beyeler C, Guttormsen S, Huwendiek S. The new final Clinical Skills examination in human medicine in Switzerland: Essential steps of exam development, implementation and evaluation, and central insights from the perspective of the national Working Group. *GMS Z Med Ausbild.* 2015;32(4):Doc40. DOI: 10.3205/zma000982
18. Schnabel KP, Lorwald AC, Beltraminelli H, Germano M, Brem BG, Wust S, Bauer D. Development and evaluation of three-dimensional transfers to depict skin conditions in simulation-based education. *GMS J Med Educ.* 2024;41(2):Doc14. DOI: 10.3205/zma001669
19. Bauer D, Lorwald AC, Wust S, Beltraminelli H, Germano M, Michel A, Schnabel KP. Development, production and evaluation of 2-dimensional transfer tattoos to simulate skin conditions in health professions education. *BMC Med Educ.* 2021;21(1):350. DOI: 10.1186/s12909-021-02763-z
20. Brem BG, Richter S, Schnabel KP. Quality Control for SP Performance Regarding Patient Portrayal in a High-Stakes OSCE. In: Association of Standardised Participants Educators (ASPE) annual Conference; June 22rd - 25th, 2014; Indianapolis (IN); 2014. p.F-P24.



21. Wind LA, Van Dalen J, Muijtjens AM, Rethans JJ. Assessing simulated patients in an educational setting: the MaSP (Maastricht Assessment of Simulated Patients). *Med Educ.* 2004;38(1):39-44. DOI: 10.1111/j.1365-2923.2004.01686.x

**Corresponding author:**

PD Dr. med. Kai P. Schnabel  
University of Bern, Institute for Medical Education,  
Mittelstr. 43, CH-3012 Bern, Switzerland  
kai.schnabel@unibe.ch

**Please cite as**

Schnabel KP, Bauer D, Schmitz FM, Hitzblech T, Brem BG. Simulated patients' role-portrayal in the clinical skills part of the Swiss federal licensing exam is of high quality and improves further over time as measured with the FAIR OSCE instrument. *GMS J Med Educ.* 2025;42(1):Doc12.  
DOI: 10.3205/zma001736, URN: urn:nbn:de:0183-zma0017366

**This article is freely available from**

<https://doi.org/10.3205/zma001736>

**Received:** 2024-02-02

**Revised:** 2024-07-12

**Accepted:** 2024-09-10

**Published:** 2025-02-17

**Copyright**

©2025 Schnabel 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/>.

# Die Qualität der Rollendarstellung von Simulationspersonen im Clinical-Skills-Teil der Eidgenössischen Prüfung der Schweiz (EPCS) ist hoch und verbessert sich stetig, gemessen mit dem FAIR-OSCE-Instrument

## Zusammenfassung

**Einleitung:** Simulationsbasierte Lehre und Prüfungen sind integraler Bestandteil der Ausbildung in den Gesundheitsberufen, wobei der Einsatz von Simulationspersonen (SP) eine weithin akzeptierte Strategie darstellt. Die Sicherstellung einer qualitativ hochwertigen SP-Rollendarstellung ist entscheidend für die Authentizität und Standardisierung von Prüfungen, insbesondere bei karriere-relevanten summativen Prüfungen wie der eidgenössischen Prüfung in Humanmedizin Clinical Skills (EPCS) der Schweiz.

Die Studie beurteilt die Qualität der SP-Rollendarstellung über mehrere aufeinanderfolgende EPCS. Wir stellten die Hypothese auf, dass sich die Qualität der Rollendarstellung mit der Zeit verbessert.

**Methoden:** In der Studie wurde das FAIR-OSCE-Instrument verwendet, um die SP-Rollendarstellung in fünf aufeinanderfolgenden EPCS von 2016 bis 2021 zu bewerten. Das Instrument, das zwischen 2011 und 2014 entwickelt wurde, umfasst vier Kategorien: Einführung, Informationsvermittlung, Rollendarstellung und weitere Elemente. Die Datenanalyse umfasste die retrospektive Überprüfung der FAIR-OSCE-Bewertungen, die Berechnung der Item-Scores und der durchschnittlichen Gesamtpunktzahl für jedes Prüfungsjahr.

**Ergebnisse:** An der Studie nahmen 37 SP-Expertinnen und Experten teil, die 1803 SP-Interaktionen mit Prüflingen an fünf Prüfungsstandorten beobachteten. Die Ergebnisse zeigten eine kontinuierliche Verbesserung der SP-Rollendarstellung über den Fünfjahreszeitraum, mit signifikanten Unterschieden zwischen 2016 und den Folgejahren. Die durchschnittlichen Gesamtwerte der SP-Rollenspielbewertungen stiegen stetig an, was auf einen positiven Trend in der SP-Leistung hindeutet.

**Diskussion:** Die Ergebnisse stützten die Hypothese einer kontinuierlichen Verbesserung der SP-Rollendarstellung innerhalb der EPCS. Die Qualität der Rollendarstellung verbesserte sich nicht nur durchweg, sondern hielt sich auch auf einem sehr hohen Niveau, wobei kein Item auf dem FAIR-OSCE-Instrument in mehr als 5% der Fälle mit „stimme nicht zu“ bewertet wurde. Dies deutet darauf hin, dass die SP-Rollendarstellungen gut mit den Fallskripten übereinstimmten, was die beabsichtigte Authentizität und Standardisierung der Bewertungen widerspiegelt.

Ein möglicher Bias bei der Bewertung der eigenen SP durch lokale SP-Experten und der ausschliessliche Fokus der Studie auf die Rollendarstellung von SP müssen hier einschränkend genannt werden. Während die Ergebnisse zum Verständnis der Effektivität von SP bei standardisierten, bestehensrelevanten klinisch-praktischen Prüfungen beitragen, wurden in der Studie andere potenzielle Varianzquellen nicht untersucht.

**Schlussfolgerung:** Zusammenfassend lässt sich sagen, dass die Studie über fünf Jahre hinweg eine kontinuierliche Verbesserung und die hohe Qualität der SP-Rollendarstellung in der EPCS zeigte. Gut ausgebildete SPs spielen, gemessen mit dem FAIR-OSCE-Instrument, eine entscheidende Rolle für die gleichbleibend hohe Qualität und Standardisierung

Kai P. Schnabel<sup>1</sup>  
Daniel Bauer<sup>1</sup>  
Felix M. Schmitz<sup>1</sup>  
Tanja Hitzblech<sup>1</sup>  
Beate G. Brem<sup>1</sup>

<sup>1</sup> Universität Bern, Institut für  
Medizinische Ausbildung,  
Bern, Schweiz

bei karriereentscheidenden klinischen Kompetenzprüfungen. Weitere Forschung könnte zusätzliche Faktoren untersuchen, die die Gesamtqualität der Prüfungen beeinflussen, und potenzielle Bias bei den Bewertungen von SP-Expertinnen und -Experten untersuchen.

**Schlüsselwörter:** OSCE, objektive strukturierte klinische Prüfung, Standardisierung, standardisierter Patient, simulierter Patient, anspruchsvolle Prüfungen, Zulassungsprüfung, Patientensimulation

## Einleitung

Simulationsbasierte Lehre und Prüfung sind für die Ausbildung und Prüfung von Gesundheitsberufen von entscheidender Bedeutung [1], [2]. Unter den verschiedenen simulationsbezogenen Optionen, wie dem Einsatz von Part-task Trainern, Full-Scale-Simulatoren und computer-gestützten Anwendungen, ist der Einsatz von Simulationspersonen (auch Simulationspatientinnen oder -patienten genannt, SP) eine bewährte Strategie in Simulationsszenarien [2], [3]. Erstmals in den 1960er Jahren eingeführt [4], [5] ist die Arbeit mit SPs in Lern- und Assessmentkontexten heutzutage von vielen medizinischen Fakultäten auf der ganzen Welt weithin akzeptiert und angewendet [6]. Objektive strukturierte klinische Prüfungen (OSCEs) [7] arbeiten regelmäßig mit SPs, in denen Medizinstudierende mit SP interagieren, die vordefinierte Patientenfälle darstellen, während die Leistungen der Studierenden bewertet werden. Die Qualität der Rollendarstellung der SPs ist von entscheidender Bedeutung, um faire und valide Prüfungen zu gewährleisten, vor allem im Hinblick auf Authentizität und Standardisierung [8]: *Unzureichend authentische* Rollendarstellung schränkt die Repräsentativität der den Prüflingen gestellten Aufgabe ein, *unzureichend standardisierte* Rollendarstellung schränkt die Vergleichbarkeit der Aufgaben und damit die Fairness in der Prüfungen ein, beides gefährdet die Validität der Prüfung [2], [9], [10]. Es ist die Aufgabe der SP-Expertinnen und Experten, dafür zu sorgen, dass die Rollendarstellung ein hohes Niveau und eine gute Balance zwischen Authentizität und Standardisierung aufweist [11], [12].

Die Schweiz hat 2011 die neue eidgenössische Prüfung Humanmedizin (EP) eingeführt [13], mit einem schriftlichen Teil und einem OSCE, den Absolventinnen und Absolventen der medizinischen Masterstudiengänge ablegen, um die Berufserlaubnis als Ärztinnen und Ärzte zu erlangen. Eine zunehmende Anzahl von Prüflingen (Absolventinnen und Absolventen des Medizinstudiums und Personen aus dem außereuropäischen Ausland, die eine Zulassung in der Schweiz anstreben) müssen in der EP bewertet werden, wobei im Jahr 2021 über tausend das Examen ablegten [14]. Es wurden verschiedene Maßnahmen ergriffen, um höchste Qualität bei der Entwicklung und Durchführung dieser Prüfung zu gewährleisten [15], [16], [17]. Die Qualität der gesamten EP spiegelt sich z. B. in Cronbachs Alpha von über 0,84 als Indikator für die Messzuverlässigkeit wider [17].

Um sicherzustellen, dass die SPs jedes einzelnen Szenarios mit ihren unterschiedlichen individuellen Hintergründen und Schauspielerfahrungen in homogenen Gruppen zusammengefasst werden, damit sie ihre jeweiligen Rollen über den gesamten Prüfungstag hinweg kontinuierlich gut darstellen, wurde ein Protokoll für die Schulung und Qualitätssicherung eingeführt. Dazu gehören 5 Tage Standardisierungstreffen auf Bundesebene, um ein gemeinsames Verständnis der Drehbücher und Szenarien und deren Umsetzung in die Praxis zu etablieren sowie Video-Material zu erstellen, das die prototypische Umsetzung jedes Szenarios darstellt (z. B. Standardisierung von Schmerzreaktionen, Auftragen von Moulagen [18], [19]). SP-Expertinnen und Experten sowie SPs der jeweiligen Prüfungsstandorte verwenden dieses Material dann pro Szenario in zwei Trainingseinheiten von jeweils 2 Stunden, um eine authentische und standardisierte Rollendarstellung während der Prüfung zu gewährleisten [17]. Während der Prüfung wird die Qualität der Rollendarstellung der SPs (und anderer Aspekte der Prüfung) durch Live-Beobachtung kontrolliert, z. B. über Einwegspiegel oder über Live-Stream in einen Kontrollraum.

Ein so komplexes wie ressourcenintensives Prüfungsformat wie das OSCE, welches sich sowohl auf Patientensicherheit auswirkt (Prüflinge, die mit falsch-positiven Ergebnissen bestehen) als auch auf die Karriere der Prüflinge (solche, die mit falsch-negativen Ergebnissen nicht bestehen) braucht notwendigerweise eine Analyse, ob die Darstellung der SP-Rollen in einem karriereentscheidenden OSCE gleichbleibend auf hohem Niveau durchgeführt wird. Das Ziel der vorliegenden Studie war es daher, die Qualität der Rollendarstellung von Simulationspersonen (SP) in klinischen High-Stakes-Prüfungen in der Schweiz über die Zeit zu bewerten. Wir stellten die Hypothese auf, dass sich die Qualität der SP-Rollendarstellung in der EPCS im Laufe der Zeit weiter verbessert.

## Methodik

Um zu beurteilen, ob die SP-Rollendarstellung in High-Stakes-Prüfungen prinzipiell und über die Zeit auf hohem Niveau erfolgt, haben wir eine Kohortenstudie mit den fünf Schweizer SP-Programmen durchgeführt, die an der Durchführung mitverantworten.

## Setting

Die EPCS wurde zum Zeitpunkt dieser Analyse an fünf Standorten (Basel, Bern, Genf, Lausanne und Zürich)



gleichzeitig auf Französisch und Deutsch über drei Tage durchgeführt, wobei jeder Durchlauf 12 Stationen umfasste, fast alle mit Einsatz von SPs (umfassende Daten zur Prüfung sind in Anhang 1 verfügbar).

## Verfahren

Um festzustellen, ob die SP-Rollendarstellung in der EPCS kontinuierlich auf hohem Niveau durchgeführt wird, wurde diese während der Durchführung der EPCS von 2016 bis 2021 analysiert, was fünf aufeinanderfolgenden Prüfungen entspricht, da die EPCS 2020 aufgrund der COVID-19-Pandemie nicht durchgeführt wurde. In diesen fünf Prüfungen wurde die klinische Kompetenz von über 5000 Prüflingen bewertet. Dabei wurden SPs eingesetzt, um insgesamt 70 parallele Durchläufe à 12 Stationen durchzuführen. Die Daten zur Rollendarstellung der SPs wurden an allen fünf Standorten während der fünf Prüfungen anhand zufälliger Beobachtungen live oder via Live-Stream unter Benutzung des FAIR-OSCE-Instruments gemessen. Die Raterinnen und Rater waren lokale SP-Expertinnen und Experten (1–5 pro Standort) und ein Beobachter mit bundesbehördlichem Mandat. Die SP-Expertinnen und Experten waren entweder Ärztinnen und Ärzte, erfahrene Gesundheitsexpertinnen und -experten oder Schauspielerinnen und Schauspieler mit 2 bis 20 Jahren Erfahrung auf diesem Gebiet und Mitglieder des eidgenössischen SP-Trainerteams.

## Instrument

Das FAIR-OSCE-Instrument wurde zwischen 2011 und 2014 von einem inter fakultären Gremium Schweizer SP-Expertinnen und Experten entwickelt, um die Rollendarstellung von SP anhand beobachtbarer, *fallunabhängiger* Kriterien zu bewerten. Es basiert auf der einschlägigen Literatur und hat einen Konsensbildungsprozess im inter fakultären Gremium durchlaufen, was seine Inhaltsvalidität belegt [15]. Das Instrument wurde zusätzlich in SP-Trainings sowie bei lokalen formativen und summativen OSCEs an der Fakultät Bern eingesetzt und hat damit seine Handhabbarkeit und Nützlichkeit bewiesen [15], [20].

Das Instrument umfasst vier Kategorien:

1. Einführung (3 Items),
2. Informationswiedergabe (7 Items),
3. Rollendarstellung (7 Items) und
4. Sonstiges (3 Items).

Die Items werden auf einer Skala mit den Punkten -1=*stimme nicht zu*, +1=*stimme zu*, +2=*stimme vollständig zu* und *nicht zutreffend* bewertet. (vgl. Anhang 2).

## Datenanalyse

Die FAIR-OSCE-Bewertungen von 2016 bis 2021 wurden retrospektiv analysiert. Die Item-Scores wurden berechnet, um die Qualität der SP-Rollendarstellung innerhalb der Prüfung eines jeden Jahres hervorzuheben und es

wurden die Mittelwerte für jede der fünf Prüfungen berechnet (2016 bis 2021, die als „nicht zutreffend“ bewerteten Items wurden ausgeschlossen), d. h. alle Beurteilungen basieren auf Daten aller Standorte, an denen die Prüfung im jeweiligen Jahr durchgeführt wurde.

Statistische Unterschiede zwischen den Jahren wurden mit einer einfaktoriellen ANOVA berechnet. Die Effektstärken wurden mit Eta-Quadrat berechnet, wobei  $\eta^2 \sim 0,01$  für kleine,  $\eta^2 \sim 0,06$  für mittlere und  $\eta^2 > 0,14$  für grosse Effekte gilt. Post-hoc-Analysen wurden mit Bonferroni-Tests für jeden möglichen Gruppenvergleich berechnet.

## Ethische Nichtbedenklichkeit

Eine Beurteilung dieser Studie durch die zuständige Ethikkommission war gemäss Schweizer Gesetzgebung (Kapitel 1, Artikel 2 des Schweizer Humanforschungsgesetzes von 2011) nicht erforderlich. Die ethischen Grundsätze gemäss der WMA-Deklaration von Helsinki und der guten wissenschaftlichen Praxis wurden durchgehend eingehalten.

## Ergebnisse

Während der fünf EPCS zwischen 2016 und 2021 beobachteten insgesamt  $N=37$  verschiedene SP-Expertinnen und Experten (ca. 2–4 pro Standort und Jahr, meist dieselben SP-Expertinnen und Experten im Laufe der Jahre) insgesamt  $N=1803$  Interaktionen zwischen SPs und Prüflingen (Basel  $n=247$ , Bern  $n=419$ , Genf  $n=357$ , Lausanne  $n=467$  und Zürich  $n=313$ ) und bewerteten die SP-Rollendarstellung.

Die Anzahl der Beobachtungen stieg von anfänglich 168 Beobachtungen im Jahr 2016 auf maximal 499 Beobachtungen im Jahr 2021 (siehe Tabelle 1).

Die Mittelwerte der SP-Rollenspiel-Bewertungen stiegen über die Jahre kontinuierlich an (siehe Tabelle 2). Die ANOVA zeigte, dass mindestens ein Anstieg über die Jahre hinweg von statistischer Signifikanz war ( $F=9,2$ ,  $p<0,001$ ;  $\eta^2=0,02$ ). Die entsprechenden Bonferroni-Post-hoc-Tests zeigten, dass die Ergebnisse von 2018 bis 2021 im Vergleich zu 2016 signifikant besser ausfielen (vgl. Tabelle 2).

## Diskussion

In Bezug auf unsere Hypothese konnten wir zeigen, dass die Qualität der SP-Rollendarstellung über die analysierten fünf Jahre kontinuierlich zunahm, signifikant war dies beim Vergleich der Daten aus 2021 mit 2016 und 2017, 2018 und 2017 mit 2016.

Damit konnten wir die anfängliche Hypothese von steigender Qualität bestätigen.

Die Qualität der Rollendarstellung verbesserte sich nicht nur kontinuierlich, sondern befand sich auch auf einem sehr hohen Niveau, wobei keine Items auf dem FAIR-OSCE-Instrument häufiger als in 5% der zufälligen Beob-

Tabelle 1: FAIR-OSCE-Antwortverteilung, exemplarisch aus der Prüfung 2021

Item	Antwortverteilung			
	Stimme vollständig zu	Stimme zu	Stimme nicht zu	Nicht beurteilbar
<i>Einleitung</i>				
<i>Erster Satz wörtlich</i>	92,2%	6,4%	0,6%	0,8%
<i>Zweiter Satz erst nach Pause/Aufforderung</i>	49,3%	13,6%	3,8%	33,3%
<i>Zweiter Satz wörtlich</i>	57,1%	8,2%	2,6%	32,1%
<i>Informationswiedergabe</i>				
<i>Korrekte Fakten</i>	86,2%	13,2%	0,4%	0,2%
<i>Korrekte Portionierung der Informationen</i>	83,8%	14,8%	1,2%	0,2%
<i>Angemessene Reaktion auf offene Fragen</i>	76,2%	13,0%	0,6%	10,2%
<i>Korrekte Weiter-/Wiedergabe von Aufforderungen/Informationen lt. Skript</i>	86,0%	12,0%	0,6%	1,4%
<i>Angemessene Improvisation</i>	67,5%	8,2%	0,2%	24,0%
<i>Keine spontanen Informationen</i>	86,6%	11,2%	1,8%	0,4%
<i>Keine spontanen Fragen</i>	93,0%	4,8%	2,0%	0,2%
<i>Rollendarstellung</i>				
<i>Korrekt verbaler Ausdruck</i>	91,2%	8,4%	0,4%	0,0%
<i>Korrekt nonverbaler Ausdruck</i>	90,0%	9,8%	0,2%	0,0%
<i>Keine nonverbalen Hinweise</i>	91,8%	7,8%	0,0%	0,4%
<i>Korrekte Darstellung von Krankheitssymptomen</i>	87,0%	12,0%	0,2%	0,8%
<i>Korrekte Darstellung von Emotionen</i>	86,2%	11,8%	0,8%	1,2%
<i>Wirkt wie ein echter Patient/Patientin</i>	91,6%	8,2%	0,0%	0,2%
<i>Bleibt in der Rolle, bis der Kandidat oder die Kandidatin den Raum verlässt.</i>	87,8%	9,2%	0,8%	2,2%
<i>Sonstiges</i>				
<i>Richtige Kleidung</i>	94,4%	2,4%	0,8%	2,4%
<i>Korrekte Moulage</i>	23,4%	0,2%	0,0%	76,4%
<i>Sorgfältiger Umgang mit Material</i>	27,1%	1,0%	0,0%	71,9%

Hinweis: gekürzte Item-Bezeichnungen

Tabelle 2: Deskriptive Statistiken zu den FAIR-OSCE-Bewertungen pro Prüfungsjahr

Jahr	Deskriptive Statistik der FAIR-OSCE-Bewertungen			
	Zahl	Mittelwert	Standardabweichung	Standardmessfehler
2016*	168	1,748	0,36	0,02
2017**	267	1,811	0,27	0,02
2018	407	1,851	0,26	0,01
2019	462	1,853	0,23	0,01
2021	499	1,869	0,19	0,01

Hinweis: \* Bonferroni-Post-hoc-Tests zeigten statistisch signifikante Unterschiede zwischen den Ergebnissen aus den Prüfungsjahren 2016 und 2018, 2019 und 2021 (alle entsprechend  $p < 0,001$ ); \*\* Bonferroni-Post-hoc-Tests ergaben einen statistisch signifikanten Unterschied zwischen den Ergebnissen aus dem Prüfungsjahr 017 und 2021 ( $p < 0,05$ ).

achtungen in allen vier Kategorien mit „*stimme nicht zu*“ bewertet wurden, was bedeutet, dass die Rollendarstellungen skriptkonform umgesetzt wurden. Beide Ergebnisse können als Indikatoren dafür gewertet werden, dass sowohl die Prüfung als auch die eingesetzten SPs von höchster Qualität sind. Das ist auch ein Indikator dafür, dass das FAIR-OSCE-Instrument in der Lage ist, die steigende Qualität der Prüfung zu messen, obwohl die Messung bereits auf einem sehr hohen Niveau begonnen hat,

wenn man den Startpunkt der EP im Jahr 2011 berücksichtigt (wo man die steilste Lernkurve erwarten würde), was das Argument der Reliabilität des Instruments untermauert.

Die Ergebnisse sind im Übrigen kein Maß für die Gesamtqualität der EPCS, sondern lediglich ein Mosaikstein, da andere potenzielle Varianzquellen, wie die Prüferinnen und Prüfer, Checklisten und weitere kontextuelle Faktoren nicht untersucht wurden und über den Rahmen dieser

Studie hinausgehen würden. Eine mögliche Quelle für einen Bias in der vorliegenden Studie liegt in der Art und Weise, wie lokale SP-Expertinnen und Experten die Qualität der Rollendarstellung ihrer SPs bewerteten, die von ihnen selbst oder ihrem Team trainiert wurden. Dies stellt jedoch einen systematischen Fehler dar, der für alle Standorte gelten würde. Das Inter-Rater-Übereinstimmung bei der Nutzung des FAIR-OSCE wurde bisher nicht untersucht und war auch nicht Gegenstand dieser Studie, sollte aber in zukünftigen Arbeiten untersucht werden. Eine parallele Bewertung mit den Authentizitäts-Items des MaSP [21] könnte die Validität von FAIR OSCE weiter untermauern. Die vorliegende Studie wurde zwar nur in einem Land und im Rahmen einer Prüfung durchgeführt, doch die Tatsache, dass sie im Rahmen einer hochstandardisierten, staatlichen Zulassungsprüfung an fünf Standorten in zwei Sprachen durchgeführt wurde, verleiht ihr einen gewissen Wert.

## Schlussfolgerung

Es zeigte sich, dass die Rollendarstellung der SPs in einer Reihe hochstandardisierter, karriereentscheidender Prüfungen von höchster Qualität war und in allen analysierten Prüfungskohorten sogar zunahm. Gut ausgebildete SPs, die mit einem geeigneten Instrument beobachtet werden, tragen somit zu einer standardisierten, qualitativ hochwertigen klinischen Kompetenzprüfung bei.

## ORCIDs der Autorinnen und Autoren

- Kai P. Schnabel: [0000-0002-6977-2717]
- Daniel Bauer: [0000-0002-3337-3327]
- Tanja Hitzblech: [0000-0003-0876-0373]
- Beate G. Brem: [0000-0002-0551-9587]

## Danksagung

Die Autorinnen und Autoren bedanken sich herzlich bei allen SP-Expertinnen und Experten, die das Instrument anwenden und die Daten während der Prüfungen zusammengetragen haben, die als Grundlage für die Analysen dienen.

## Interessenkonflikt

Die Autorinnen und Autoren erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

## Anhänge

Verfügbar unter <https://doi.org/10.3205/zma001736>

1. Anhang\_1.pdf (114 KB)  
Umfassende Beschreibung des Teils Klinische Fertigkeiten der Eidgenössischen Prüfung in Humanmedizin Clinical Skills (EPCS)
2. Anhang\_2.pdf (272 KB)  
Deutsche Fassung des FAIR OSCE-Instruments in seiner derzeitigen Form

## Literatur

1. Bradley P. The history of simulation in medical education and possible future directions. *Med Educ.* 2006;40(3):254-62. DOI: 10.1111/j.1365-2929.2006.02394.x
2. Bauer D, Lahner FM, Huwendiek S, Schmitz FM, Guttormsen S. An overview of and approach to selecting appropriate patient representations in teaching and summative assessment in medical education. *Swiss Med Wkly.* 2020;150:w20382. DOI: 10.4414/sm.w.2020.20382
3. Nestel D, Kneebone R. Perspective: authentic patient perspectives in simulations for procedural and surgical skills. *Acad Med.* 2010;85(5):889-893. DOI: 10.1097/ACM.0b013e3181d749ac
4. Barrows HS, Abrahamson S. The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. *J Med Educ.* 1964;39:802-5.
5. Barrows HS. Simulated patients in medical teaching. *Can Med Assoc J.* 1968;98(14):674-676.
6. Nestel D, Bearman M. *Simulated Patient Methodology.* Malden (MA); Oxford (UK): West Sussex UK: John Wiley & Sons Ltd.; 2015. DOI: 10.1002/9781118760673
7. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
8. Smith CM, Gliva-Mcconvey G, Chapin A. Training SPs for Authentic Role Portrayal. In: Gliva-McConvey G, Nicholas CF, Clark L, editors. *Comprehensive Healthcare Simulation: Implementing Best Practices in Standardized Patient Methodology Comprehensive Healthcare Simulation.* Cham: Springer; 2020. DOI: 10.1007/978-3-030-43826-5
9. Haladyna TM, Downing SM. Construct-Irrelevant Variance in High-Stakes Testing. *Educ Measurement.* 2004;23(1):17-27. DOI: 10.1111/j.1745-3992.2004.tb00149.x
10. Smith CM, O'Byrne CO, Nestel D. Simulated patient methodology and assessment. In: Nestel D, Bearman M, editors. *Simulated Patient Methodology.* Malden (MA); Oxford (UK): West Sussex UK: John Wiley & Sons Ltd.; 2015. p.85-92. DOI: 10.1002/9781118760673.ch12
11. Lewis KL, Bohnert CA, Gammon WL, Hölzer H, Lyman L, Smith C, Thompson TM, Wallace A, Gliva-Mcconvey G. The Association of Standardized Patient Educators (ASPE) Standards of Best Practice (SOBP). *Adv Simul (Lond).* 2017;2:10. DOI: 10.1186/s41077-017-0043-4
12. Brem B, Steinweg B. Prüfungen mit SPs. In: Peters T, Thrien C, editor. *Simulationspatienten Handbuch für die Aus- und Weiterbildung in medizinischen und Gesundheitsberufen.* 1. Auflage. Göttingen: Hogrefe Verlag; 2018.

13. Guttormsen S, Beyeler C, Bonvin R, Feller S, Schirlo C, Schnabel K, Schurter T, Berendonk C. The new licencing examination for human medicine: from concept to implementation. *Swiss Med Wkly*. 2013;143:w13897. DOI: 10.4414/smw.2013.13897
14. Bundesamt für Gesundheit. Statistiken Ärztinnen/Ärzte. Bern: Bundesamt für Gesundheit; 2022. Zugänglich unter/available from: [https://www.bag.admin.ch/dam/bag/de/dokumente/berufe-gesundheitswesen/medizinalberufe/statistiken/med/aerztinnen-aerzte-2021.pdf.download.pdf/Aerztinnen\\_und\\_Aerzte\\_2021\\_DE.pdf](https://www.bag.admin.ch/dam/bag/de/dokumente/berufe-gesundheitswesen/medizinalberufe/statistiken/med/aerztinnen-aerzte-2021.pdf.download.pdf/Aerztinnen_und_Aerzte_2021_DE.pdf)
15. Brem B, Richter S, Schnabel KP. FAIR\_OSCE - Feedbackstructure for assessment of interactive roleplay in Objective Structured Clinical Exams. In: 4th Swiss Conference on Standardized Patients and Simulation in Health Care. Bern, 10-12 September 2014. SC2-5. Zugänglich unter/available from: <https://with-simulation.ch/programm-2014/>
16. Bachmann C, Kropf R, Biller S, Schnabel KP, Junod Perron N, Monti M, Berendonk C, Huwendiek S, Breckwoldt J. Development and national consensus finding on patient-centred high stakes communication skills assessments for the Swiss Federal Licensing Examination in Medicine. *Patient Educ Couns*. 2021;104(7):1765-1772. DOI: 10.1016/j.pec.2020.12.003
17. Berendonk C, Schirlo C, Balestra G, Bonvin R, Feller S, Huber P, Junger E, Monti M, Schnabel K, Beyeler C, Guttormsen S, Huwendiek S. The new final Clinical Skills examination in human medicine in Switzerland: Essential steps of exam development, implementation and evaluation, and central insights from the perspective of the national Working Group. *GMS Z Med Ausbild*. 2015;32(4):Doc40. DOI: 10.3205/zma000982
18. Schnabel KP, Lorwald AC, Beltraminelli H, Germano M, Brem BG, Wust S, Bauer D. Development and evaluation of three-dimensional transfers to depict skin conditions in simulation-based education. *GMS J Med Educ*. 2024;41(2):Doc14. DOI: 10.3205/zma001669
19. Bauer D, Lorwald AC, Wust S, Beltraminelli H, Germano M, Michel A, Schnabel KP. Development, production and evaluation of 2-dimensional transfer tattoos to simulate skin conditions in health professions education. *BMC Med Educ*. 2021;21(1):350. DOI: 10.1186/s12909-021-02763-z
20. Brem BG, Richter S, Schnabel KP. Quality Control for SP Performance Regarding Patient Portrayal in a High-Stakes OSCE. In: Association of Standardised Participants Educators (ASPE) annual Conference; June 22rd – 25th, 2014; Indianapolis (IN); 2014. p.F-P24.
21. Wind LA, Van Dalen J, Muijtjens AM, Rethans JJ. Assessing simulated patients in an educational setting: the MaSP (Maastricht Assessment of Simulated Patients). *Med Educ*. 2004;38(1):39-44. DOI: 10.1111/j.1365-2923.2004.01686.x

**Korrespondenzadresse:**

PD Dr. med. Kai P. Schnabel  
 Universität Bern, Institut für Medizinische Ausbildung,  
 Mittelstr. 43, CH-3012 Bern, Schweiz  
[kai.schnabel@unibe.ch](mailto:kai.schnabel@unibe.ch)

**Bitte zitieren als**

*Schnabel KP, Bauer D, Schmitz FM, Hitzblech T, Brem BG. Simulated patients' role-portrayal in the clinical skills part of the Swiss federal licensing exam is of high quality and improves further over time as measured with the FAIR OSCE instrument. GMS J Med Educ. 2025;42(1):Doc12.*

DOI: 10.3205/zma001736, URN: urn:nbn:de:0183-zma0017366

**Artikel online frei zugänglich unter**

<https://doi.org/10.3205/zma001736>

**Eingereicht:** 02.02.2024

**Überarbeitet:** 12.07.2024

**Angenommen:** 10.09.2024

**Veröffentlicht:** 17.02.2025

**Copyright**

©2025 Schnabel 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/>.