

RESEARCH

Open Access



Exploring the perception of medical students and lecturers on the consequential validity of medical long case

Jacob Kumakech^{1*}, Ian Guyton Munabi², Aloysius Gonzaga Mubuuke³ and Sarah Kiguli⁴

Abstract

Introduction The long case is a historical clinical assessment method for medical students but is being phased out due to low reliability arising from the difficulty in standardization of the assessment. However, it's an educationally valuable assessment, whose consequential validity has got little attention in the literature.

Objective The aim of the study was to explore the consequential validity of the long case as perceived by medical students and lecturers.

Methodology The study followed an interpretivist approach utilizing qualitative methods of data collection. Data was collected through Focus Group Discussions with medical students ($n=6$) as well as Key Informant Interviews with lecturers ($n=5$). The qualitative data was analyzed by deductive qualitative analysis utilizing Cook and Lineberry framework for evaluating the educational impact of an assessment.

Results & discussion The study found that the long case is an educationally valuable assessment method because it is authentic and provides a holistic approach to clinical assessment. The assessment promotes professional identity development as students assume the role of a physician. Though reliability was reported as the main weakness of the long case, participants were strongly in favor of the long case especially for formative assessment. It was also noted that awarding marks in the formative assessment serves as a major incentive for students to do more long cases. Though, the study reported patients being uncooperative to students sometimes, it lacked the voices of the patients yet they are a major player during the long case assessment.

Conclusion The long case is an educationally valuable assessment for formative evaluation of medical students and offers great potential for professional development.

Keywords Long case, Assessment, Perceptions, Consequential validity, Education impact of an assessment

*Correspondence:

Jacob Kumakech
jkumakech30@gmail.com

¹School of Medicine, Department of Paediatrics & Child Health, Makerere University, Kampala, Uganda

²School of Biomedical Sciences, Department of Anatomy, Makerere University, Kampala, Uganda

³School of Medicine, Department of Radiology, Makerere University, Kampala, Uganda

⁴School of Medicine, Department of Pediatrics & Child Health, Makerere University, Kampala, Uganda



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

The traditional long case has become unpopular in medical schools in the Global North in favor of Objective Structured Clinical Exam (OSCE) and Work Base Assessment such as Mini Clinical Evaluation Exercise (Mini-Cex) which are believed to yield more reliable and easily defensible results [1–3]. Yet, in our previous investigation on the experiences of medical students and lecturers with the long case, we reported that the assessment method is still utilized in both formative and summative assessment of medical students in the Global South (low-resource settings) [4]. This form of assessment exposes students to patients, and provides an opportunity for students to be observed and assessed in the performance of the real tasks of a physician [4, 5]. The long case provides a comprehensive assessment of clinical skills unlike OSCEs, which encourage a reductionist approach by splitting clinical tasks into constituent parts and testing them separately [2, 6]. This is in agreement with the notion that the whole is more than the sum of its parts [7], as artificially breaking down a one-hour long case clinical encounter into several bites may not be sufficient to assess the complex clinical skills and competence of a medical trainee. Therefore, the long case assesses an integrated approach to clinical medicine which enables students to see patients wholly rather than as sum of their diseased organ systems. The long case is thus an educationally valuable assessment with high face validity [8]. It is authentic, located in a clinical platform, and provides a holistic assessment of clinical skills, doctor patient relationship, professionalism, communication and teamwork [9–12]. In keeping with constructivism, the long case is a meaningful assessment which reflects the practice of the profession or discipline being assessed [13] providing an assessment of the ability to apply knowledge in the context in which one would work after graduation.

However, there are two main arguments against the long case that is its inter-rater and inter-case reliability. The low inter-rater reliability is due to the lack of standardization of the long case [4]. The examiners often use flexible, subjective, and global judgment, with a lack of clarity in marking scale. Thus a student examined via the long case is often scored differently by different examiners due to variation in leniency or stringency of the examiner [14–16]. Additionally, inter-case reliability is a major concern due to variation in diagnostic complexities, as some cases are generally more complex than others. Thus, a single long case cannot assess the full breadth of skills or sample the curriculum widely [17, 18], and it is difficult to generalize from one long case encounter about student ability in other cases [19]. However, reliability is a major concern mainly when a student performance is judged based on a single long case as is often done summative assessment. In formative assessment

however, a student makes several long cases over time and encounters varying degrees of case complexities on different topics and is often assessed by different lecturers [4] which improves reliability.

Despite the concerns over reliability, the long case continues to be a cornerstone in clinical assessment in the Global South probably because of its educational impact. The construct of educational impact, also referred to as consequential validity, has two dimensions [20]. Firstly, the assessment needs to have accurate descriptions in order to guide interpretation. Secondly, and more important to this study, is that the consequential validity appraise the potential and actual consequences of using an assessment method including both intended and unintended outcome extending to even social consequences of an assessment [20–22]. However, there is a sizeable research gap on the long case with no published literature on its consequential validity in the Global South. Therefore, the aim of this study was to use a qualitative research method to explore consequential validity of the long case as perceived by medical students and lecturers to offer guidance to educators on how to optimize, effectively utilize and exploit the potentials of the long case in clinical assessment while providing quality medical education and patient care.

Methodology

Study design

An interpretivist approach was taken in this exploratory qualitative study. The data was collected through focus group discussions and interviews, and analyzed by deductive thematic analysis.

Study setting

The study was conducted at Makerere University School of Medicine in Uganda, where the Bachelor of Medicine and Bachelor of Surgery (MBChB) degree is a five-year program. The first two years are for pre-clinical training and the last three years are dedicated to clinical clerkship. The medical students complete a junior and senior clerkship in Pediatric, Internal Medicine, General Surgery and Obstetrics and Gynecology in the third- and fifth-years respectively. The fourth year is dedicated to clinical experience in various medical and surgical specialties. The long case is used in formative assessment during clerkship in all four major disciplines. In this formative assessment, medical students present their long case during ward rounds on a daily basis. The students assume the role of a doctor who watches over patients assigned under his/her care and presents the patient for discussion with the lecturers, residents and their peers. The student does a minimum of 10 long cases over the clerkship period of 7-weeks for which the student is formatively assessed. In internal Medicine Clerkship however,

the long case is also used in summative assessment during which a student presents only one case at the end of the semester for which they are questioned and assessed. For each long case, the medical students interact with patients to take a medical history, perform physical examinations, formulate a diagnosis and treatment plan. The student then presents their long cases to a lecturer or a resident in the presence of the patient and peers. This is then followed by questioning and discussion of the case.

Study participants

The study participants ($n=42$) were third- and fifth-year medical students and internal medicine lecturers. At the time of data collection, the total numbers of third and fifth year medical students were 191 and 188 respectively and the department of internal medicine had 42 physicians/lecturers. The study participants were recruited by purposive sampling and the sample size was determined by data saturation point. The third-year students ($n=19$) recruited were those who had completed long case assessments in junior clerkship while the fifth-year students ($n=18$) were recruited on basis of having completed the long case in both junior and senior clerkships. Of the third-year students recruited, there were ten male and nine female while the fifth-year participant comprised of eight male and ten female. The lecturers in the internal medicine department ($n=5$) were recruited to participate because they assess the long case during both formative and summative assessment. Out of the five lecturers who participated, three were female and two male. The students were invited through their class representative and a brief recruitment presentation was made at the study site, while the lecturers were approached via email and telephone invitation.

Data collection

Ethical approval to conduct the study was obtained from Makerere University School of medicine research and ethics committee with ethics ID Mak-SOM-REC-2022-524. Informed consent was obtained from all the study participants.

Data were collected through Focus Group Discussions (FGDs) and Key Informant Interviews (KIIs). The FGD and KII guides were developed and piloted prior to the study (supplementary file 1 and 2). It was a semi structured guide which provided a set of questions while allowing freedom to ask in any order, following tangents or seeking clarification of previous answers or elaboration of responses.

The study was conducted in May & June 2023 during recess term when student have just completed academic year. A total of 36 medical students participated in FGDs, reflecting on their experiences with the long case. Six face to face FGDs were conducted, three for junior clerks

and three for senior clerks. Each FGD comprised of 5–7 participants with balanced male and female gender representation. A research assistant proficient in qualitative research methods, without prior knowledge about the long case and having no relationship to the students facilitated the FGDs. The discussions lasted between 55 min and 1 h 10 min and were audio recorded, and transcribed verbatim. Data saturation was achieved by the fifth FGD, at which point no additional new information was shared.

Semi-structured KIIs were used to collect data from Internal Medicine faculty. Five KIIs were conducted, and data saturation was achieved by the fourth interview, at which point no new information was shared. The Principal Investigator, a graduate from the same medical school more than 4 years before the study, having no current ties with internal medicine faculty conducted the KIIs via Zoom. Each interview lasted between 25 and 50 min and was recorded and transcribed verbatim. The data collected were securely stored on a hard drive and Google Drive with password protection to prevent unauthorized access.

Data analysis

Data analysis was done through deductive thematic analysis method. The use of predetermined categories to analyze data makes deductive analysis more structured, systematic and efficient [23]. It begins with a research question that clearly connects to the theory or framework to be used in analysis. The researchers then operationalize the theory or framework by either generating sensitizing constructs, or key concepts from them that are revised throughout analysis [23, 24]. It is noteworthy that deductive analysis may limit the discovery of unexpected patterns and promote confirmation bias, as the analysis is confined to predefined themes [24]. In our data analysis, Cook and Lineberry framework for organizing data on the impact of education assessment [25] was used. The framework proposes that the education impact of an assessment should be explored by considering the impact on various stakeholders such as students, educators, schools, or the end target of practice (e.g., patients or health care systems); and the downstream impact of classification like the different score cut points. The impact can result from the assessment activities or use of the score; they can be intended or unintended, beneficial or harmful [25]. The framework enabled systematic analysis of the consequence of the long case assessment in medical education. During data analysis, we defined the research question then applied the predetermined coding framework, then coding segments of text according to the predefined themes. The data was coded in multiple iterations, resulting in rich, nuanced understanding of the phenomenon under investigation. Finally, the coded data were reviewed and organized to

Table 1 Table of results showing themes, subthemes and codes generated from the study

| Themes | Subthemes | Codes |
|---|---|---|
| Theme 1: Impact on students. | More genuine learning Professional growth Communication and team work | Superior to reading medicine from books fosters Professional identity formation Inter-professional communication, confidence, and team spirit |
| Theme 2: Social impact | Impact on lecturers Impact on patients Impact on the hospital | Overwhelmed by The student number, challenge to balance teaching with provision of clinical care. Patients get tired of being clerked, refusal to consent to students Congestion on the hospital ward |
| Theme 3: Impact of the assessment process | Authenticity Foster expertise Lack of standardization | It is real life and holistic Promotes growth from novice to expert Some cases are more complex than others, difficult to standardize |
| Theme 4: Impact of marks | Marks as an incentive Subjectivity of the marks Marks or no marks | Marks encourages students to go to wards Marks depend on the assessor, some lecturers are mean while others are generous with mark Awarding marks in formative assessment, long case is not fit for summative assessment. |

ensure they align with the thematic framework, allowing researchers to refine or consolidate key themes as necessary. Throughout the analysis, the researchers stayed true to the framework while remaining open to unexpected findings to limit the loss of subtle nuances or other possible understanding.

Results

Four themes emerged from the study (Table 1) namely: impact on students, social impact which encompassed lecturers, patients and the hospital, impact of the assessment process, and finally the impact of marks.

Theme 1: impact on students

Students believe that long case facilitate their learning of clinical medicine more effectively than if they are to read text-book describing it as providing more genuine and diverse definition of diseases.

Diseases don't read books; you will read in the book that malaria presents like this but you reach the hospital and this person is presenting differently and the diagnosis is malaria (FGD 1 participant 3).

In addition to providing genuine learning, the long case fosters professionalism as medical students begin to model professional behaviors through interaction with the patient resulting in professional identity formation.

It also gives you the feel that I am a doctor, even if you're not licensed. You also feel like you're useful, you're not just going to come on ward but you can report about and actually follow up a patient (FGD 1 Participant 1).

While the lecturers maintain that it fosters practical application and assessment of medical ethics, a key domain in health care.

When you come to the ward, you should know that you are dealing with human being but in the skills lab, you are not going to develop that skill (KII 5).

Relatedly, both lecturers and students reported that the long case improves communication skills, inspires confidence, and instills inter-professional practice and teamwork spirits.

How do you communicate with patients and in an interdisciplinary team, to your colleagues; how do you express yourself to communicate a point (KII 3). I think it also improves on your public speaking skills the more you present these cases to your colleagues and your senior. You learn more how to talk to other people with confidence (FGD 2 Participant 6).

Theme 2: social impact

Both lecturers and students believe that the number of students on the wards is prohibiting, making it hard for every student to get as many opportunities to present as possible.

Clinical training is like nurturing somebody to behave or conduct themselves in a certain way. So if the numbers are many the impacts per person reduce and the quality goes down (KII 5). Mulago has a heavy workload with many patients but few consultants who may not get a chance to listen to every student; they may listen to like two to save time to see all the patients FGD 3 Participant 4).

Related, the lecturers find it challenging to balance addressing the urgent concerns of patient and providing good long case and learning experience to students within the limited time.

The challenge is that as a lecturer you're tasked with providing clinical care and at the same time teach-

ing. I think it's impossible to do both extremely well for the reason that there's a limitation of time. You wouldn't see every patient in extreme detail to provide an experience for the learner to learn every detail. If we all do that then the mortality rates in the ward would be high (KII 1).

The students reported that some patients are often not willing to share their medical information or allow students to examine them, and often patients get tired of being clerked by many students.

I personally sometimes face a problem of a non-cooperative patient. At some point patients are irritated, maybe because we are many who are asking them the same questions all the time, every day (FGD 4 Participant 4).

Additionally, the wards are congested and there is no barrier separating one patient from the next. This undermines confidentiality of patient information when students present their long cases on such wards.

There are so many students, the ward is full, and we are even taking away the oxygen of patients (FGD 1 Participant 6).

Sometimes these wards are very congested, as you're presenting to this consultant who wants you to be audible enough you be like this patient is HIV positive. On saying that, the patient looks down because the two immediate neighbors now know (FGD 1 Participant 4).

Theme 3: impact of the assessment process

In terms of the impact of the long case as an assessment process, lecturers acknowledged that the long case is difficult to standardize as the different cases vary in complexities.

If a student is seeing somebody who has malaria versus one who is seeing patient with say an autoimmune illness, obviously the one with auto immune is much more complex (KII 1).

Yet, they also emphasized that the authenticity of the assessment process is a real strength.

They are practical, hands on, face to face, you can interact, discuss as opposed to when it's just for example an MCQ (KII 2).

Finally, in addition to describing the long case as authentic, real life, and hands on, students also mentioned that the long case process help them meet their learning goals.

I think it really helps us meet our learning objectives because after the presentation you discuss this case with the doctor (FGD 4 Participant 2).

Theme 4: impact of marks

Both students and lecturers believe that the mark (grade) is an incentive that drives students to continue going to the wards.

I wouldn't even go to the hospital if there is no mark. I can remain in my room and read medicine (FGD 1 Participant 3).

Awarding a mark provide functions as an incentive to the Learner. There has to be some mechanism of the learner feeling that I'm getting better; a measure of gradual improvement between both the learner and the lecturer (KII 1).

However, some students believe that removing the marks would take away the pressure of hunting for marks and focus on learning.

I think if they remove away marks, it will help focus on learning other than chasing the marks because that's the first thing students chase for. So it takes away that pressure and it will make you free to learn with ease (FGD 1 Participant 1).

Therefore, to exploit the benefits as an incentive while at the same reducing pressure to hunt for the marks, students believe that the long case should be used solely as formative assessment but not in summative evaluation. And that their performance in formative evaluation should be graded and used to compute their score/achievement in clinical clerkship at the end of the academic semester.

If I have been doing it at least once in a week, then in the whole rotation I've done a minimum of ten. So, I do that continuously instead of being forced to cram in the last one week to present one case which may even turn out to be a mess. They can say that we are taking the best ten, to compute your passing score; so that even while rotating, I know that I'm supposed to prepare myself well so that I make a nice presentation (FGD 6 Participant 2).

Both students and lecturers however, consider it a very subjective assessment. The marks are dependent on who assesses you as some lecturers may be biased toward a student, others are either mean or generous with marks depending on your luck. This is even worse in the single case presented for summative assessment.

The assessment is obviously subjective; I can't say that it is absolutely objective (KII 1).

It's so subjective, you are already told of those lecturers who give marks in the range of 50 to 60%. So in your presentation, you're going to put in as much effort but you know end result is 50/60; they are so malignant (FGD 2 Participant 7).

Discussion

This study followed an interpretivist approach to examine the consequential validity of the medical long case as perceived by medical students and lecturers utilizing Cook and Lineberry framework for evaluating education impact of an assessment. The study highlighted the benefit of authentic learning and assessment experience with much emphasis on multiple exposures as in formative assessment. It also provided an insight into the role of the long case in fostering professional growth and professional identity formation among medical students.

The study findings revealed that the long case fosters professional identity formation among medical students, as the students expressed getting a realistic glimpse into the experience of being a doctor. Importantly, the participants believe it enables the students to develop several aspects of medical professionalism such as communication, ethical skill and inter-professional practice and team spirit which are vital for a medical practitioner. This finding suggests a strong connection between taking ownership of patient care and forming a professional identity. It reverberates with research findings that students who engage deeply with simulated patients tend to reflect on their future roles as physicians and the responsibilities they entail [26–28]. The ownership of real patients would foster an even more genuine sense of professional identity development compared to simulated patients. Thus highlighting how ownership of real patients in the context of long case enables students to reflect on their role as physician. The finding aligns with the constructivist perspective that meaningful assessment should mirror professional practice [12, 29]. It is therefore vital for medical students to be exposed to the art being a doctor during training, something that would strengthen their conviction to the profession but also expose them to real challenge as they navigate the complexity of clinical encounter during the long case.

The study underscored the role of an authentic assessment method and learning environment. It highlighted the fact that long case provides unparalleled holistic real-life experience that successfully addresses student learning goals and objectives while allowing lecturers to evaluate all the competencies required of a physician. This finding is similar to previous studies on the long case which reported benefits such a holistic approach [12]. Authentic learning and assessment offer learners

practical applications of their knowledge that are relevant to real-life situations. By linking these authentic tasks to their existing knowledge, students can build deeper, more lasting connections.

From the study, awarding marks to students during long case is a significant motivator to indulge in the activities. Students are more diligent in their preparation when they know that each long case presentation impacts their final assessment [30, 31]. This underscores the vital role of assessment in driving learning as students feel incentivized to work harder. Additionally, marks provide a mechanism of monitoring the students' progress over time. However, subjectivity of the long case was a major challenge reported especially when student is assessed based on a single case as it is the norm in summative assessment. Subjectivity was reportedly due to examiner bias and difference in the stringency or leniency of assessors [15]. Subjectivity has been reported in several previous studies as a cause of low inter-reliability of the long case [14, 16, 30]. However, the students preferred to do several long cases throughout the semester as it is in formative assessment, as a way of mitigating this weakness. This is consistent with previous studies that emphasized increasing the number of student-patient encounters is the single most important step in rehabilitating the long case [32, 33], as it would give students exposure to diverse patient conditions enabling them to refine their skills. This continuous learning process, characterized by ongoing experiences and the gradual accumulation of knowledge and skills, supports the development of understanding [34]. This study therefore strongly recommends against assessing student's clinical skills based on a single long case as it is often the practice in summative assessment. Therefore, the alternative is to assess several long cases and incorporate the evidence in portfolio as in formative assessment throughout the clerkship. This approach could address the concern about reliance on subjective evaluation for high-stakes grading.

The findings also highlighted the importance of balancing the demand of patient care and teaching/assessing students. Relatedly, it divulged the unwillingness of some patients to consent to students or allow students to participate in their management. These findings present an important ethical dilemma as patient autonomy and beneficence is important yet at the same time there is a responsibility to provide training needs of the medical students. The unwillingness of the patients to consent to students could probably be due to lack of privacy on the ward, too many students clerking the same patient, or patients' lack of understanding about the role of students on the ward. Mwaka, Taremwa et al. (2022) reported similarly about patients discomfort with the involvement of medical students in their care [35]. This presents an important area for further studies and raises critical

questions for educators. As students go into the clinical areas, there is need to balance patient experience with the student, while providing timely patient care in the face of the need to teach. This finding also presents one critical limitation the study; that is lack of patient voices on the subject under investigation. Thus future study is necessary to understand patient's perspective and concerns in the context of the long case and students' presence in clinical environment.

Conclusion

This research into the consequential validity of the medical long case among medical students and lecturers provides strong evidence that the long case is effective in driving student learning. It provides a holistic and authentic experience that mirrors the tasks of physician, imparting a strong sense of professional identity among the medical students as it does not only assess clinical knowledge and skills, but also professionalism, doctor patient relationship, communication and team work. These are very important findings that underscore the importance of a patient-centered approach to assessment as in the long case.

However, the long case is noted to present the same challenges of low reliability due to lack of standardization. The low reliability is however an issue mainly when used in summative assessment. Therefore, it is important to assess several long cases throughout the clerkship and incorporate the evidence in a portfolio as in formative assessment. The study therefore strongly recommend against the uses of long case in Summative assessment to address the concerns about reliance on subjective evaluation for high-stakes grading.

Finally, this study revealed two ethical dilemmas with the long case and student presence in patient environment: Firstly patients' hesitancy to consent to students, a finding that needs further research as the reasons behind the hesitancy and the experience of patients were not explored due to lack of patient voice in this study. Secondly lecturer's need to balance effective patient care and providing good long case experience to the students. Thus, effective strategies need to be put in place to enable lecturers meet both demands effectively.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12909-025-07055-4>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

I acknowledge the support from Paper with Purpose (PwP) funded by British academy. Papers with Purpose is a structured writing development and support program designed to empower early career Health Professions

Education (HPE) researchers from across the African continent to produce impactful publications. The established HPE researchers from the University of cape town and Oxford university who facilitated the PwP workshops provided valuable insight, critique, comments and suggestions for improvement to the work in this publication.

Author contributions

JK contributed to the conception and design of the study and the acquisition, analysis and interpretation of the data. He made the first draft of the work. IMG an M.A contributed to the analysis and interpretation of the data. Made major corrections of the first draft of the manuscript. SK; Made major corrections to the first draft and made the final corrections to the submitted version of the work. All authors read and approved the submitted version. The authors all agreed to be both personally accountable for their contribution and to ensure that questions related to the accuracy of integrity of any part of the work, even one in which he/she was not personally involved, are appropriately investigated and resolved and that the resolution is documented in the literature.

Funding

This research was supported by the Forgarty International Centre of the National Institute of Health under award number 1R25TW011213. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of Health.

Data availability

The data supporting the results of this article can also be accessed from the Makerere University repository, Perceptions of Medical Students and Lecturers of the Long Case Practices as Formative Assessment in Internal Medicine Clerkship at Makerere University repository, 2024, identifier <http://hdl.handle.net/10570/13032>; additional data and the raw data are available with the Principal Investigator.

Declarations

Ethical approval

This research was conducted in compliance with the Declaration of Helsinki with the Ethical approval to conduct the study obtained from the Makerere University School of Medicine Research and Ethics Committee; ethics ID Mak-SOMREC-2022-524. Informed consent was obtained from all the study participants. Strict confidentiality was maintained and participants were identified with unique identification number.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Clinical trial number

not applicable.

Received: 21 November 2024 / Accepted: 25 March 2025

Published online: 22 April 2025

References

1. Khan H. OSCEs are outdated: clinical skills assessment should be centred around workplace-based assessments (WPBAS) to put the 'art' back into medicine. *MedEdPublish* 2017, 6.
2. Al-Wardy NM. Assessment methods in undergraduate medical education. *Sultan Qaboos Univ Med J*. 2010;10(2):203.
3. Pell G, Roberts T. Setting standards for student assessment. *Int J Res Method Educ*. 2006;29(1):91–103.
4. Kumakech J, Munabi IG, Mubuuu AG, Kiguli S. Experiences of medical students and faculty regarding the use of long case as a formative assessment method at a tertiary care teaching hospital in a low resource setting: a qualitative study. *BMC Med Educ* 2024, 24.

5. Dare AJ, Cardinal A, Kolbe J, Bagg W. What can the history tell Us?? An argument for observed history-taking in the trainee intern long case assessment. *Clin Correspondence* 2008;1.
6. Tey C, Chiavaroli N, Ryan A. Perceived educational impact of the medical student long case: a qualitative study. *BMC Med Educ*. 2020;20:1–9.
7. Upton J, Janeka I, Ferraro N. The whole is more than the sum of its parts: Aristotle, metaphysical. *J Craniofac Surg*. 2014;25(1):59–63.
8. Shrivastava SR, Shrivastava PS. Scope of Long-Case examination as an assessment tool among medical students and the way forward. *J Integr Health Sci*. 2021;9(1):42–3.
9. Zeller A, Battegay M, Gyr N, Battegay E. Evaluation of unstructured medical school examinations: prospective observational study. *Swiss Med Wkly*. 2003;133(1112):184–184.
10. Wass V, Van Der Vleuten C. The long case. *Med Educ*. 2004;38(11):1176–80.
11. Olson CJ, Rolfe I, Hensley. The effect of a structured question grid on the validity and perceived fairness of a medical long case assessment. *Med Educ*. 2000;34(1):46–52.
12. Masih CS, Benson C. The long case as a formative assessment Tool–Views of medical students. *Ulster Med J*. 2019;88(2):124.
13. Reeves TC, Okey JR. Alternative assessment for constructivist learning environments. *Constructivist Learn Environments: Case Stud Instructional Des*. 1996;191:202.
14. Leichner P, Sisler GC, Harper D. A study of the reliability of the clinical oral examination in psychiatry. *Can J Psychiatry*. 1984;29(5):394–7.
15. McManus IC, Thompson M, Mollon J. Assessment of examiner leniency and stringency ('hawk-dove effect') in the MRCP (UK) clinical examination (PACES) using multi-facet Rasch modelling. *BMC Med Educ*. 2006;6:1–22.
16. Wilson G, Lever R, Harden RM, Robertson J, MacRitchie J. Examination of clinical examiners. *Lancet*. 1969;293(7584):37–40.
17. Benning T, Broadhurst M. The long case is dead–long live the long case: loss of the MRCPsych long case and holism in psychiatry. *Psychiatr Bull*. 2007;31(12):441–2.
18. Ponnamperna GG, Karunathilake IM, McAleer S, Davis MH. The long case and its modifications: a literature review. *Med Educ*. 2009;43(10):936–41.
19. Burn W, Brittlebank A. The long case: the case against its revival: commentary on... the long case. *Psychiatrist*. 2013;37(12):382–3.
20. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol*. 1995;50(9):741.
21. Iliescu D, Greiff S. On consequential validity. In.: Hogrefe Publishing; 2021.
22. Shepard LA. The centrality of test use and consequences for test validity. *Educational Measurement: Issues Pract*. 1997;16(2):5–24.
23. Gilgun JF. Deductive qualitative analysis and grounded theory: sensitizing concepts and hypothesis-testing. *SAGE Handb Curr Developments Grounded Theory* 2019:107–22.
24. Fife ST, Gossner JD. Deductive qualitative analysis: evaluating, expanding, and refining theory. *Int J Qualitative Methods*. 2024;23:16094069241244856.
25. Cook DA, Lineberry M. Consequences validity evidence: evaluating the impact of educational assessments. *Acad Med*. 2016;91(6):785–95.
26. Wyatt TR. Investigating the meaning of patient ownership: an exploratory study of a commonly used phrase within an internal medicine department. *The medical/health Humanities-Politics, programs, and pedagogies*. edn.: Springer; 2022. pp. 241–50.
27. Tien L, Wyatt TR, Tews M, Kleinheksel A. Simulation as a tool to promote professional identity formation and patient ownership in medical students. *Simul Gaming*. 2019;50(6):711–24.
28. Mukhalalati B, Aly A, Yakti O, Elshami S, Daud A, Awaisu A, Sethi A, El-Awaisi A, Stewart D, Abu-Hijleh MF, et al. Examining the perception of undergraduate health professional students of their learning environment, learning experience and professional identity development: a mixed-methods study. *BMC Med Educ*. 2024;24(1):886.
29. Scholtz A. An analysis of the impact of an authentic assessment strategy on student performance in a technology-mediated constructivist classroom: a study revisited. *Int J Educ Dev Using ICT*. 2007;3(4):42–53.
30. Tey C, Chiavaroli N, Ryan A. Perceived educational impact of the medical student long case: a qualitative study. *BMC Med Educ*. 2020;20(1):1–9.
31. Teoh NC, Bowden FJ. The case for resurrecting the long case. *BMJ*. 2008;336(7655):1250–1250.
32. Kroboth FJ, Hanusa BH, Parker S, Coulehan JL, Kapoor WN, Brown FH, Karpf M, Levey GS. The inter-rater reliability and internal consistency of a clinical evaluation exercise. *J Gen Intern Med*. 1992;7(2):174–9.
33. Norcini JJ. The death of the long case? *Bmj* 2002, 324(7334):408–9.
34. Jaworski B. Constructivism and teaching-the socio-cultural context. *Retrieved October* 1996, 26:2011.
35. Mwaka AD, Taremwa S, Adoch W, Achan J, Ainembabazi P, Walego G, Ntayi ML, Bongomin F, Ibingira CB. Patients' attitudes towards involvement of medical students in their care at university teaching hospitals of three public universities in Uganda: a cross sectional study. *BMC Med Educ*. 2022;22(1):1–16.

Publisher's note

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