PRACTICAL TIPS

**Twelve tips for introducing the concept of validity argument in assessment to novice medical teachers in a workshop**

**[version 2]**

Hosam Eldeen Gasmalla¹,4, Majed Wadi²,5, Mohamed H. Taha³

¹Al-Neelain University  
²Al-Neelain University  
³Medical Education Department  
⁴Medical Education Department  
⁵University of Sharjah

**Abstract**

This article was migrated. The article was marked as recommended.

Background: Misconceptions have been observed in the application of validity by faculty and in the reporting of validity in a significant amount of published work in the field of students' assessment. As a result, actions concerning the dissemination of information about the concept of validity in relation to assessments, especially among novice medical teachers, is needed.

Aim: This work aims to guide how the concept of validity argument in assessments is delivered to novice medical teachers in a workshop.

Methods: Critical reflection and a careful review of relevant literature were used to develop these tips.

Results and Conclusion: Twelve tips were introduced to support instructors conducting workshops on introducing the concept of validity, especially to novice medical teachers.

**Keywords**

Validity evidence, Assessment, Measurement

---

**Open Peer Review**

**Migrated Content**

"Migrated Content" refers to articles submitted to and published in the publication before moving to the current platform. These articles are static and cannot be updated.

| Version | View | View |
|---------|------|------|
| version 2 | 21 Sep 2021 |  |
| version 1 | 25 Mar 2021 |  |

1. **Barbara Jennings**, Norwich Medical School  
2. **Ken Masters**, Sultan Qaboos University  
3. **Julie Hunt**, Lincoln Memorial University

Any reports and responses or comments on the article can be found at the end of the article.
**Introduction**

The interpretation of the term validity has changed over time. In the 1920s and up to the 1950s, it referred to the degree to which a test measures what it is supposed to measure. During that timeframe, emphasis was placed on the idea of validity being the character of the test; at that time, criterion validity and content validity were the two known types of validity. When the correlation of test scores to a given criterion is determined, this is referred to as criterion validity (Colliver, Conlee and Verhulst, 2012), while content validity refers to the relation between the content of the assessment tool and the measured construct (Artino et al., 2014). Thus, tests were defined as being either valid or not. In the mid-1950s, a third type of validity emerged: construct validity. This was due to the struggle to define the criterion/reference standard of some traits, such as clinical reasoning, using criterion and content validity because, by that time, the concept of validity was associated with an explicit purpose. In the 1980s, the contemporary notion of validity evolved, and it was seen as a unitary concept evolving around the construct (AERA, APA and NCME, 1999). The concept of validity has continued to evolve; it is now focused on the suitability and appropriateness of the interpretations made from assessment scores (Oermann and Gaberson, 2019).

Misconceptions and malpractices have been noted in the reporting of validity in the published work that focuses on different types of assessment methods (Gasmalla and Tahir, 2020). The practice of validating the interpretation of test scores has been described as being less than optimal (St-Onge and Young, 2015). Investigators often apply deficient approaches to the validity argument (Cook et al., 2015). Furthermore, there is uncertainty among researchers regarding the type and amount of validity evidence that must be collected and presented to sufficiently support their inferences, and the identification of the applied validity frameworks is either incomplete or absent. Sometimes, outdated frameworks are used (Cook et al., 2014). With various conceptualizations of validity, researchers do not adhere to a unified approach when reporting the validity of an assessment (Young et al., 2018).

Since misconceptions have been observed regarding how the concept of validity is understood, action is needed to clarify the best ways to disseminate information about this concept. This work aims to provide guidance on educating novice medical teachers on the concept of validity with regard to assessments via a workshop. The characteristics of a novice medical teacher have been described by Persky and Robinson (2017), and these include the tendency to abide by rules, guidelines, and plans. They need to obtain information in order to learn, and they spend time on the recall process.

This work is especially suited for faculty development units or experts who choose to conduct workshops as a training method. It is based on the author’s experience and those cited in the literature.

**Tip 1. Consider the number and the needs of participants**

It is imperative to limit the number of workshop participants to a range of 5 to 15. This will allow for better engagement in the discussion, and it will allow the participants to benefit from the exercises offered during the workshop (McLeod et al., 2010). However, restricting the number of participants to 15 may not always be a feasible option. In such cases, the instructor/facilitator of the workshop is advised to pay more attention to strategies that can improve the engagement of the participants, such as adding more interactivity, emphasizing dialogue-based styles, integrating storytelling, and adopting tech-enabled approaches. Previous studies have emphasized the benefits of these approaches (Kinsella, Mahon and Lillis, 2017). After determining the number of participants, it is crucial to conduct a training needs assessment (TNA). A TNA helps in the assessment of the faculty’s current level of knowledge about the concept of validity, and it compares that level of knowledge to the required standard. Rather than simply assuming that all faculty need training or the same level of training on the concept of validity, organizers should use this tool to determine training needs. This point is further explored in the next tip.

Many factors affect the structure and duration of the workshop, such as the number of participants and the TNA. A suggested timetable is attached in the appendix (Table 1).

**Tip 2. Explore the workshop participants’ prior knowledge**

There are many validity frameworks. Traditionally, validity has been classified into content, criterion, and construct validity (Goodwin and Leech, 2003). Validity is also viewed as being a unitary concept (Downing, 2003), and there is Kane’s framework for validation (Cook et al., 2015). Since novice medical teachers may have varying degrees of prior knowledge about the concept of validity, it is crucial to assess their understanding of validity before starting the training.

The exploration of prior knowledge is based on constructivism learning theory. This theory posits that new knowledge and comprehension develop when an individual’s existing understanding is built upon (Badyal and Singh, 2017). Learners construct knowledge based on their experiences, and they assimilate, accommodate, and adapt knowledge to develop a new understanding (Torre et al., 2006). The learning process entails the role of critical reflection, which is based...
on prior experiences, to construct meaning (Cirocki and Farrell, 2017). The different validity frameworks could be contributing to the confusion about validity, especially with the ongoing discussion that challenges the concept of construct validity (Collier, Conlee and Verhulst, 2012). Thus, it is useful to start by exploring the participants’ knowledge of validity. Prior knowledge is explored to enrich the discussion, and the subsequent debate and disagreement among the workshop participants will help ensure that their minds are open and receptive to what comes next.

**Tip 3. Consider the historical evolution of the concept**

After exploring the different views on validity, it is time to provide an explanation of it by discussing the historical evolution of the concept. This will help later during the explanation of the concept of validity. For simplification, this information can be presented in the chronological order of the evolution of the concept. For example, the presentation can begin from the era of content and criterion validity (1920s-1950s), followed by the introduction of construct validity in the mid-1950s (Cronbach and Meehl, 1955), which led to the era of the tripartite concept, and finally the introduction of the unitary concept of validity (AERA, APA and NCME, 1999). During this step, it is imperative to explain why the tripartite concept has failed to grasp the full magnitude of validity and that the definition of validity has evolved from a focus on the validity of the test to the validity of its use for a particular purpose and, eventually, the validity of the interpretations drawn from the test scores. This will smoothly guide the participants toward the concept of a construct since all the aspects of validity are centered around it.

**Tip 4. Explain the concept of a construct**

Since the contemporary concept of validity has been seen as a unitary concept that evolves around the construct, it is important to discuss what is meant by a construct. A construct is a characteristic that a test aims to measure (AERA, APA and NCME, 1999). It is an intangible learner characteristic that cannot be detected directly. Examples include the ability to solve clinical problems or the ability to recall. In this step, the workshop participants are asked to take their time explaining this concept, as it will later help them comprehend the importance of defining the construct and linking all the other validation activities to it. Thus, the purpose of this step is to explain how other types of validity contribute in one way or another to construct validity.

**Tip 5. Start backwards: Discuss the threats to validity**

At this point in the workshop, it is not yet time to introduce the five sources of validity evidence. From our experience, it is useful to present a scenario: you are in a faculty board meeting or department meeting, and one of your colleagues has presented the test score of a student who has failed. In your opinion, what could have possibly gone wrong? This question will generate a discussion about threats to validity. You will not have to mention them; you will only need to categorize the examples the participants provide into construct-overrepresentation and construct-irrelevant variance (Downing and Haladyna, 2004) Thus, the purpose of this step is to introduce the workshop participants to the threats to validity.

**Tip 6. Ask the workshop participants to define the impact of the threats on inference**

At this point in the workshop, you may lead the discussion into the effects of threats on validity. The overall process requires the participants to present examples of threats to validity, describe the effects, and finally, suggest the actions that need to be taken to eliminate those threats (see Tip 7).

This tip (Tip 6) will usually be interwoven with the previous tip and the following tips. The following diagram (Figure 1) explains the process for Tips 5, 6, and 7.

**Tip 7. Ask the participants to provide examples of how to eliminate the threats to validity**

Downing has provided a comprehensive guide describing the threats to validity, including examples of how to eliminate them (Downing and Haladyna, 2004). These threats are divided into two categories: construct under-representation (CU) and construct-irrelevant variance (CIV). An example of the former category (CU) is the use of few test items, while the latter category includes any approach that leads to bias in the interpretation of test results (e.g., flawed items). Asking the participants to mention examples of actions taken to eliminate threats to validity will prompt them to think about providing evidence for those actions, which will include evidence of validity. This will pave the way for the introduction of the five sources of evidence of validity. Sources of validity evidence can be gathered according to the following categories: content, response process, internal structure, relationship to other variables, and consequences (Downing, 2003). It is useful to conduct an exercise in which every possible piece of evidence is categorized into the five sources; such an exercise will maximize the participants’ engagement.

**Tip 8. End at the beginning: Explain the unitary concept of validity**

Now that the participants are familiar with the history of the changes in the concept of validity, the notion of construct validity, validity threats, and the five sources of validity evidence, it is time to tie it all together and solidify the unitary concept of validity. By this point in the workshop, you will find that your job has become easy.
Tip 9. Ask the participants to compare the tripartite concept of validity to the current concept of validity

By now, some of the participants will still be centered on the old concept of validity (that is based on the tripartite notion). (Goodwin and Leech, 2003) provided a useful way to aid in the participants’ transition from the tripartite concept to the unitary concept by comparing the types of validity evidence between the old concept and the current concept. The purpose of this step is to ease the transition for the participants who are still in favor of the traditional concept.

Tip 10. Address the misconceptions

The comparison between the tripartite concept and the current concept will very likely reveal the number of misconceptions that are still out there. Examples of misconceptions may include the use of the phrase validity of a test instead of validity evidence that supports the interpretation or the idea that any amount (or type) of validity evidence is enough to build the case rather than a focus on the importance of each piece of evidence in relation to the construct and the assessment tool. Furthermore, some may think that validation is just a collection of evidence rather than a whole process that includes the identification of the construct and an evaluation of the evidence. One of the false impressions is that validity is a dichotomy (it is either valid or not valid) as opposed to a continuum. Some estimate reliability and rely on it as the sole source of validity, while others think that after establishing a validity argument for a test, the test is valid for any further/different uses. Gasmalla and Tahir (2020) elaborated on these misconceptions.

Tip 11. Ask the participants to validate the test scores from their own previously developed tests

At this point, it is time for the participants to apply what they have learned. Thus, it is necessary to ask them to bring any previous tests they have developed along with the test scores. They can work in groups. Their validation activity may raise some questions that were addressed previously, or it may reveal unnoticed misconceptions. As the instructor, you must be patient since the concept of validity is complex, ever-evolving, and the subject of continuous debate.

Tip 12. Post-credit scene: Promote a culture of quality assurance

By the end of the workshop, participants may question the need to discuss such a complex concept. In this step, a culture of quality assurance should be promoted. Assuming the workshop has gone as planned, the participants will find themselves thinking about quality assurance at every step in their career in education. In fact, the promotion of quality assurance should be highlighted throughout the workshop.
Conclusion
This work aims to provide guidance about educating novice medical teachers about the concept of validity argument with regard to assessments using a workshop. Critical reflection and a careful review of the literature were applied to develop these tips. The tips focused on the history of the evolution of the concept of validity, the notion of construct validity, validity threats, and the five sources of validity evidence; they also addressed the common misconceptions that accompany the application of validity. These tips were arranged in such a way to help the workshop participants understand the concept of validity.

Take Home Messages
- The interpretation of the term validity has changed over time.
- Misconceptions and malpractices have been noted in the reporting of validity.
- Introducing the concept of validity requires a systematic approach. This approach includes a consideration of the historical evolution of the concept, a discussion on the threats to validity, an explanation of the unitary concept of validity, a comparison of the tripartite concept of validity with the current concept of validity, addressing misconceptions, and the promotion of a culture of quality assurance.
- The backbone of this approach involves conducting a TNA followed by an effective faculty development program.

Notes On Contributors
Hosam Eldeen Elsadig Gasmalla is an assistant professor of human anatomy and histology in the faculty of Medicine, Al Neelain University and a researcher in the Education Development Center, Sudan International University. ORCID ID: https://orcid.org/0000-0003-2590-8587

Majed Mohamed Saleh Wadi is a senior lecturer of Medical Education, College of Medicine, University of Qassim, KSA. He is the Head of Medical Education Department. Head of the Assessment Committee. ORCID ID: https://orcid.org/0000-0002-8117-770X

Mohamed Hassan Taha is a visiting Assistant Professor of Medical Education at the College of Medicine and Medical Education Centre, University of Sharjah, UAE. He is the chair of the Curriculum Committee as well as the Faculty Development Committee. ORCID ID: https://orcid.org/0000-0003-0808-5590

Appendices

| Time/duration | Session/concept                                      | Mode of delivery | Applied tip                                                                 |
|---------------|------------------------------------------------------|------------------|----------------------------------------------------------------------------|
| 8:00–8:30     | Welcome, icebreakers                                |                  |                                                                            |
| 8:30–8:45     | Pertest                                              |                  |                                                                            |
| 8:45–10:00 AM | Validity in students’ assessments: What do we know about it? Historical background | Interactive session Debate | Tip #2 Explore the workshop participants’ prior knowledge  
Tip #3 Consider the historical evolution of the concept |
| 10:00–10:30 AM| Coffee break                                         |                  |                                                                            |
| 10:30–11:30 AM| Threats to validity and how to eliminate them        | Group work Group presentation | Tip #5 Start backwards: Discuss the threats to validity  
Tip #6 Ask the workshop participants to define the impact of the threats on inference  
Tip #7 Start backwards: Discuss the threats to validity |
Table 1: Continued

| Time/duration   | Session/concept                                                                 | Mode of delivery          | Applied tip                                                                 |
|----------------|--------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------|
| 11:30–12:30 PM | Validity concepts<br>Validity as a unitary concept<br>Previous and current concepts of validity: A comparison | Interactive session Debate | Tip #8 End at the beginning: Explain the unitary concept of validity<br>Tip #9 Ask the participants to compare the tripartite concept of validity to the current concept of validity |
| 12:30–13:00 PM | Coffee break                                                                    |                            |                                                                              |
| 13:00–14:00 PM | Addressing the misconceptions                                                   | Interactive session       | Tip #10 Address the misconceptions                                           |
| 14:00–15:00 PM | Exercises                                                                       | Group work Group presentation | Tip #11 Ask the participants to validate the test scores from their own previously developed tests<br>Tip #12 Post-credit scene: Promote a culture of quality assurance |
| 15:00–15:15 PM | Post-test                                                                        |                            |                                                                              |
| 15:15–15:30 PM | Wrap up/summary and take-home message                                           |                            |                                                                              |

Declarations
The author has declared that there are no conflicts of interest.

Ethics Statement
These guidelines/twelve tips are developed entirely by the authors based on their experience.

External Funding
This article has not had any External Funding.

Supplementary Files
Supplementary file 1 - Version Two - Twelve tips for introducing the concept of validity in assessment to novice medical teachers in a workshop.docx

Acknowledgments
The authors are grateful to Prof Mohamed Nour-El-Din Saleh, Assessment Unit, College of Medicine, Qassim University and Dr Husameldin Khalafalla from College of Medicine, the University of Jazan, Saudi Arabia, for their generous support and valuable comments during the preparation of this manuscript.

Figure 1 was developed by the authors.

Bibliography/References
AERA, APA and NCME (1999) Standards for Educational and Psychological Testing. American Educational Research Association. Available at: Reference Source (Accessed: 16 December 2020).
Artingo, A. R. Jr., La Rochelle, J. S., Dezee, K. J. and Gehlbach, H. (2014) Developing questionnaires for educational research: AMEE Guide No. 87. Medical Teacher. 36(6), pp. 463–74. Reference Source
Badyal, D. K. and Singh, T. (2017) Learning Theories: The Basics to Learn in Medical Education. Int J Appl Basic Med Res. 7(Suppl 1), pp. 51–53. Reference Source
Cirocki, A. and Farrell, T. S. (2017) Reflective practice for professional development of TESOL practitioners. The European Journal of Applied Linguistics and TEFL. 6(2), pp. 5–23.
Collier, J. A., Conlee, M. J. and Verhulst, S. J. (2012) From test validity to construct validity ... and back?. Medical Education. 46(4), pp. 366–71. Reference Source
Cook, D. A., Brydges, R., Ginsburg, S. and Hatala, R. (2015) A contemporary approach to validity arguments: a practical guide to Kane’s framework. Medical Education. 49(6), pp. 560–75. Reference Source
Cook, D. A., Zendejas, B., Hamstra, S. J., Hatala, R., et al. (2014) What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. Adv Health Sci Educ Theory Pract. 19(2), pp. 233–50. Reference Source
Cronbach, L. J. and Meehl, P. E. (1955) Construct validity in psychological tests. Psychological Bulletin. 52(4), p. 281.
Downing, S. M. (2003) *Validity: on the meaningful interpretation of assessment data*. Medical Education. 37(9), pp. 830–837.

Downing, S. M. and Haladyna, T. M. (2004) *Validity threats: overcoming interference with proposed interpretations of assessment data*. Medical Education. 38(3), pp. 327–33.

Gasmalla, H. E. E. and Tahir, M. E. (2020) *The validity argument: Addressing the misconceptions*. Medical Teacher. pp. 1–8.

Goodwin, L. D. and Leech, N. L. (2003) *The meaning of validity in the new Standards For Educational and Psychological Testing: Implications for measurement courses*. Measurement and Evaluation in Counseling and Development. 36(3), pp. 181–191.

Kinsella, G. K., Mahon, C. and Lillis, S. (2017) *Facilitating Active Engagement of the University Student in a Large-Group Setting Using Group Work Activities*. Journal of College Science Teaching. 46(6).

McLeod, P., Steinert, Y., Boudreau, D., Snell, L., et al. (2010) *Twelve tips for conducting a medical education journal club*. Medical Teacher. 32(5), pp. 368–70.

Oermann, M. H. and Gaberson, K. B. (2019) *Evaluation and Testing in Nursing Education, Sixth Edition*. Springer Publishing Company.

Persky, A. M. and Robinson, J. D. (2017) *Moving from novice to expertise and its implications for instruction*. American Journal of Pharmaceutical Education. 81(9).

St-Onge, C. and Young, M. (2015) *Evolving conceptualisations of validity: impact on the process and outcome of assessment*. Medical Education. 49(6), pp. 548–550.

Torre, D. M., Daley, B. J., Sebastian, J. L. and Elnicki, D. M. (2006) *Overview of current learning theories for medical educators*. The American Journal of Medicine. 119(10), pp. 903–907.

Young, M., St-Onge, C., Xiao, J., Vachon Lachiver, E., et al. (2018) *Characterizing the literature on validity and assessment in medical education: a bibliometric study*. Perspect Med Educ. 7(3), pp. 182–191.
Open Peer Review

Migrated Content

Version 2

Reviewer Report 12 October 2021

https://doi.org/10.21956/mep.20213.r31297

© 2021 Hunt J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Julie Hunt
Lincoln Memorial University

This review has been migrated. The reviewer awarded 4 stars out of 5

In this 12 tips article, the authors discuss how to structure and run a workshop teaching about current validity theory for educators. I appreciate the authors' method of introducing threats to validity and having educators work with those before grappling with validity frameworks, sources of evidence, etc. Overall this is an important topic for educators, as all of us make assessments of our students with varying implications for their progression, and these instructions would help educators or administrators to set up a workshop for faculty.

Competing Interests: No conflicts of interest were disclosed.

Reviewer Report 22 September 2021

https://doi.org/10.21956/mep.20213.r31296

© 2021 Masters K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ken Masters
Sultan Qaboos University

This review has been migrated. The reviewer awarded 4 stars out of 5

The authors have addressed most of my concerns about Version 1 of the article, and have produced an improved Version 2. I think that the article will be very useful for those who wish to run workshops on the
topic. Well done.

**Competing Interests:** No conflicts of interest were disclosed.

---

**Version 1**

**Reviewer Report 13 June 2021**

https://doi.org/10.21956/mep.19018.r27297

© 2021 Masters K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ken Masters
Sultan Qaboos University

This review has been migrated. The reviewer awarded 3 stars out of 5

An interesting 12-tips paper on introducing the concept of validity in assessment to novice medical teachers in a workshop. Overall, the paper is a nice blend of theoretical background and practical tips. Many important ideas are raised. There are, however, some gaps, and I would like to see these addressed:

• Tip 1: Although restricting the number of people to 15 is generally a good idea in workshops, in most organisations, this is not feasible. The needs of the institution are so great, and running the workshop multiple times is usually not possible. So, although it would go against the advice of the authors, it would help if they could address the question of how to mitigate the problems if presenters were forced to run the workshop for 25-30 participants. (For many of these presenters, “It cannot be done” is not an option).

• Similarly, the paper should give an idea of a timetable. While I understand that this may depend upon the results of the needs analysis, a rough guide (perhaps as an appendix) would be useful.

• Figure 1: There is a typo (“elminat”).

• Tip 7: It would be useful here to give a summary of Downing’s guide.

• Tip 8 and elsewhere: As there are several crucial terms and definitions, I would like to see the paper summarise them, so that they could be used by presenters to be given to the participants.

• Related to the point above: One of the roles of a 12-tips paper is that it can stand alone and assist the novice. The readers can trust that the authors have done all the background reading, and now present some form of summarised, off-the-shelf package that can be used. In many instances, this kind of practical assistance is offered in the paper, but there are gaps. Perhaps a way to test this is to give the paper to a person who knows little about the topic, ask them to use the paper to design such a workshop, and have them explain how things would be done. In this way, the gaps would become more evident and could be addressed. So, I would like to see Version 2 of the paper in which these issues are addressed.

**Competing Interests:** No conflicts of interest were disclosed.
Barbara Jennings  
Norwich Medical School

This review has been migrated. The reviewer awarded 3 stars out of 5

Thank you to the authors for an interesting and thought-provoking article about the important topic of validity as a concept in assessment. The literature cited spans many decades & includes recent articles that consider problematic misconceptions - and throughout the article the idea of an evolving definition for validity (within medical education and assessment) is revisited. I think there are a few problems with the article in its current form however, and these could be considered by the authors for a revised version; 1. The field is jargon-heavy and the introduction (and some tips in the main body) does (do) not have enough explanations in plain English for the general medical educator; or indeed, explicit descriptions of the concepts raised. For example, what are the precise threats to validity raised by Downing and others? 2. The tips are presented with a notional workshop in mind. I am unclear about which participants the authors had in mind for such a workshop – could the authors be explicit about this or perhaps consider a more generalisable model. An assessment team may be expected to know a great deal about the topic but many medical educators and even module leads may not have any preconceived ideas or enough background knowledge to warrant working through the particular steps (though would need to know about the broad concept around validity and quality assurance). This takes me to concern number 3. 3. Why focus on a history of the field at all in a workshop (tips 2/3)? Beginning with a unitary concept of validity from the start of any training seems the best approach to me – can we make progress if we constantly justify the evolution within a field? 4. A minor problem is typos and syntax issues to correct in abstract, main body and the figure. I think the article could be more successful if it presented “twelve tips about what medical educators should know about the concept of validity” – rather than a design for a very particular workshop that may not have wide-application?

**Competing Interests:** No conflicts of interest were disclosed.