Considering Students' Perceptions of the Face Validity in Formative Quizzes: A Proposal for a Mixed-Methods Process for Analyzing Opinions on Content Value

2025 Assessment Institute in Indianapolis

Quinton Quagliano, M.S., C.S.P. Psychology Department

Table of Contents

1 Introduction

1.1 Follow Along

1.2 Motivation

- *Valid* classroom assessment is **hard**
- Student perceptions on evaluations are **unclear**
- Student's have opinions on the *instructor*, less on *content*
- How can I bring students into the **process** of how I "grade" them

1.3 Value of Student Buy-in

- Students may see assessment as accountability at best, irrelevant at worst (Brown & Hirschfeld, 2008; Fletcher et al., 2012)
- Trust is important so they can learn from assessment (Carless, 2009; Chu et al., 2014; Leighton & Bustos Gómez, 2018)
- The right, little "moves" can help make that trust more tangible (Felten et al., 2023)

2 My Implementation

2.1 Assessment at Regular Interval

- I already had consistent weekly quizzes as a mainstay in my courses
- Associated with good effects on retention (Haigh, 2007; Thomas et al., 2020)
- When corrective feedback is fast, supports better self-correction (Marcell, 2008)
- My existing structure was ripe for addition

2.2 Embed in the Assessment

- 10-question multiple-choice quiz, open-note, relatively low stakes
- Embed in 1 - 2 quant/qual questions to assess “accuracy” of quiz
- Aim to capture an immediate “state”-like feeling, knee-jerk reaction (similar to Spielberger (2012))

2.3 Format of the Question

- Research on face validity from the student perspective is sparse at times
- Does show it’s importance in maintaining relevance to the student (Hollis-Sawyer & Sawyer, 2008; Sato & Ikeda, 2015)
- Some examples:
  - “Rate how close this quiz was to what we learned in class last week (1 - 5)”
  - “Do you feel like this content was covered in class? (yes / no)”
  - “Explain whether you feel this was representative of what you know (open)”
  - “What other topic were you expecting to show up that didn’t (open)”

2.4 Debrief and Analysis

- If easily accessible through LMS, may display bar graphs on board for numerical questions
- Look for common vocabulary in the open responses
- Use the insights from the qualitative to follow-up on the quantitative

3 Advice for Your Implementation

3.1 Things to Avoid

- Creating culture of arguing for points
- Students aren’t SMEs, you don’t need to treat them like one!
- Only using quant or qual - use both!
- Trying to do too much!

3.2 References

Allaire, J., & Dervieux, C. (2025). *Quarto: R interface to quarto markdown publishing system*. https://github.com/quarto-dev/quarto-r

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, *15*(1), 3–17. https://doi.org/10.1080/09695940701876003

Carless, D. (2009). Trust, distrust and their impact on assessment reform. *Assessment & Evaluation in Higher Education*, *34*(1), 79–89. https://doi.org/10.1080/02602930801895786

Chu, M.-W., Guo, Q., & Leighton, J. P. (2014). Students' interpersonal trust and attitudes towards standardised tests: Exploring affective variables related to student assessment. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 167–192. https://doi.org/10.1080/0969594X.2013.844094

Felten, P., Forsyth, R., & Sutherland, K. A. (2023). Building trust in the classroom: A conceptual model for teachers, scholars, and academic developers in higher education. *Teaching and Learning Inquiry*, *11*. https://doi.org/10.20343/teachlearninqu.11.20

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, *64*(1), 119–133. https://doi.org/10.1007/s10734-011-9484-1

Haigh, M. (2007). Sustaining learning through assessment: An evaluation of the value of a weekly class quiz. *Assessment & Evaluation in Higher Education*, *32*(4), 457–474. https://doi.org/10.1080/02602930600898593

Hollis-Sawyer, L. A., & Sawyer, T. P. (2008). Potential stereotype threat and face validity effects on cognitive-based test performance in the classroom. *Educational Psychology*, *28*(3), 291–304. https://doi.org/10.1080/01443410701532313

Leighton, J. P., & Bustos Gómez, M. C. (2018). A pedagogical alliance for trust, wellbeing and the identification of errors for learning and formative assessment. *Educational Psychology*, *38*(3), 381–406. https://doi.org/10.1080/01443410.2017.1390073

Marcell, M. (2008). Effectiveness of regular online quizzing in increasing class participation and preparation. *International Journal for the Scholarship of Teaching and Learning*, *2*(1). https://doi.org/10.20429/ijsotl.2008.020107

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sato, T., & Ikeda, N. (2015). Test-taker perception of what test items measure: A potential impact of face validity on student learning. *Language Testing in Asia*, *5*(1), 10. https://doi.org/10.1186/s40468-015-0019-z

Spielberger, C. D. (2012). *State-trait anxiety inventory for adults*. https://doi.org/10.1037/t06496-000

Thomas, A. K., Smith, A. M., Kamal, K., & Gordon, L. T. (2020). Should you use frequent quizzing in your college course? Giving up 20 minutes of lecture time may pay off. *Journal of Applied Research in Memory and Cognition*, *9*(1), 83–95. https://doi.org/10.1037/h0101845

Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in R. In V. Stodden,

F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research*. Chapman; Hall/CRC.

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC. https://yihui.org/knitr/

Xie, Y. (2025). *Knitr: A general-purpose package for dynamic report generation in r*. https://yihui.org/knitr/