JMBE Journal of Microbiology & Biology Education

There Is More than Multiple Choice: Crowd-Sourced Assessment Tips for Online, Hybrid, and Face-to-Face Environments

Jennifer C. Drew,[a] Neal Grandgenett,[b] Elizabeth A. Dinsdale,[c] Luis E. Vázquez Quiñones,[d] Sebastian Galindo,[e] William R. Morgan,[f] Mark Pauley,[g] Anne Rosenwald,[h] Eric W. Triplett,[a] William Tapprich,[i] and Adam J. Kleinschmit[j]

[a]*Department of Microbiology and Cell Science, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, USA*
[b]*Department of Teacher Education, University of Nebraska at Omaha, Omaha, Nebraska, USA*
[c]*College of Science and Engineering, Flinders University, Bedford Park, South Australia, Australia*
[d]*Division of Science and Technology, Universidad Ana G. Méndez–Cupey Campus, San Juan, Puerto Rico*
[e]*Department of Agricultural Education and Communication, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, USA*
[f]*Department of Biology, College of Wooster, Wooster, Ohio, USA*
[g]*Division of Undergraduate Education, National Science Foundation, Alexandria, Virginia, USA*
[h]*Department of Biology, Georgetown University, Washington, DC, USA*
[i]*Department of Biology, University of Nebraska at Omaha, Omaha, Nebraska, USA*
[j]*Department of Natural and Applied Sciences, University of Dubuque, Dubuque, Iowa, USA*

KEYWORDS assessment, distance learning, online, undergraduate biology education, Network for the Integration of Bioinformatics in the Life Sciences (NIBLSE), bioinformatics, Bloom's taxonomy, reliability, validity

INTRODUCTION

Life science educators have responded to the recommendations in *Vision and Change* (1–3) to improve undergraduate education by developing new courses and programs in bioinformatics, biomechanics, systems biology, and other emerging, interdisciplinary fields. These initiatives reflect that quantitative measurements, advanced technology, and data science are becoming increasingly important in biology. However, many instructors are struggling to implement these changes due to a lack of preparation time, resources, and assessments aligned with updated instructional competencies and learning resources (4). Further, recent attention to issues of justice, equity, inclusion, and diversity requires faculty to employ creative ways to fairly assess all learners (5).

Supported by new technologies, higher education has experienced a rapid evolution in available teaching modalities (e.g., blended, HyFlex, synchronous, asynchronous online) (6). The widespread adoption of online teaching, most recently spurred by the COVID-19 pandemic, will remain an important part of undergraduate education; therefore, student assessments will need to be innovatively structured for online, face-to-face, and blended environments (7). There is thus an immediate need for assessments that are valid, reliable, and flexible in new learning environments.

Challenges associated with curriculum development and adapting to new teaching modalities present an opportunity for instructors to implement the core *Vision and Change* action items by aligning assessments to learning goals and integrating multiple forms of assessment to track student learning (1–3). To help instructors implement new assessment tools or refresh current assessment strategies, we have prepared this summary of assessment types that are appropriate for multiple learning environments. Although bioinformatics, an emerging interdisciplinary field, is the theme, the assessment types are widely applicable to other fields. The provided example assessments are mapped to bioinformatics core competencies (8) to model how

Citation Drew JC, Grandgenett N, Dinsdale EA, Vázquez Quiñones LE, Galindo S, Morgan WR, Pauley M, Rosenwald A, Triplett EW, Tapprich W, Kleinschmit AJ. 2021. There is more than multiple choice: crowd-sourced assessment tips for online, hybrid, and face-to-face environments. J Microbiol Biol Educ 22:e00205-21. https://doi.org/10.1128/jmbe.00205-21.
Address correspondence to Department of Natural and Applied Sciences, University of Dubuque, Dubuque, Iowa, USA. E-mail: AKleinschmit@dbq.edu.
Received: 18 February 2021, Accepted: 12 August 2021, Published: 30 September 2021

NIBLSE Core Competencies

C1. Explain role of computation and data mining in hypothesis driven biology
C2. Summarize key concepts and applications
C3. Apply statistical concepts
C4. Use bioinformatics tools to examine complex problems
C5. Find, retrieve, and organize biological data
C6. Explore/model biological interactions or networks
C7. Use command-line tools and write scripts
C8. Describe/manage bio data types
C9. Interpret ethical, legal, medical, and social implications

FIG 1. The nine NIBLSE bioinformatics core competencies for undergraduate biologists. See reference 8 for the full description of each competency.

assessments can align to learning outcomes that include both concepts and skills.

PROCEDURE

The Network for the Integration of Bioinformatics in Life Sciences Education (NIBLSE) is an NSF-funded Research Coordination Network for Undergraduate Biology Education (9, 10). NIBLSE has established a set of bioinformatics core competencies for undergraduate biologists (Fig. 1) and is working to provide vetted bioinformatics learning resources (4, 10, 11). The NIBLSE Assessment Validation Committee (AVC) compiles, reviews, and aligns assessments to these core competencies. Although not an exhaustive list, the summary presented here describes 10 assessment types used regularly in undergraduate teaching by NIBLSE members and other bioinformatics faculty (Appendix 1 in the supplemental material). All question types have been used in face-to-face, blended, and online modalities and were submitted by NIBLSE steering committee members and instructors who completed a survey (4). Here, we provide a brief summary of each type and discuss trade-offs, along with providing a crowd-sourced exemplar of a bioinformatics-based assessment aligned to a student learning outcome and a core competency for an undergraduate course (Appendix 1 in the supplemental material).

CONCLUSION

Within the context of effective assessment, it is important to consider two features: validity and reliability (12). Considering these two features here is timely, as a recent analysis of the quality of bioinformatics assessments found that <1% of studies assessing student learning gains mentioned the use of both validity and reliability measures (13).

Validity relates to actually measuring what one seeks to measure. For example, if bioinformatics is the stated focus of a test, the test would not be valid if it only addressed basic biology concepts. There are various ways to measure validity and different types of validity, such as "content validity" (an assessment measures the targeted content of a field of knowledge adequately and sufficiently), "construct validity" (an assessment measures the intended knowledge or skills), and "concurrent validity" (an assessment that correlates well with a previously validated instrument) (14). A simple initial step to help ensure content validity is to have colleagues in the same field review and critique an assessment. Construct validity can be tested with a small group of novice students verbally describing their interpretation of assessment questions.

Reliability relates to how consistently a test produces the same scores when taken by similarly prepared students. There are various ways to demonstrate reliability, such as "test-retest," "internal consistency," and "parallel forms" (15). Typically, reliability is demonstrated by giving a particular test two or more times, while looking at how consistent the results of a test are when students have not had additional learning interventions. Useful statistical procedures for examining reliability are provided by the Web Center for Social Research Methods (16).

Importantly, assessment questions should strive to discriminate between higher and lower levels of cognitive learning according to Bloom's taxonomy (17). It is also important to separate out those questions that contribute effectively to the overall assessment and those that lower overall assessment reliability. A common strategy that is often built into Learning Management System (LMS) environments is the item discrimination index (Fig. 2) (18). This is a correlation coefficient (point-biserial based) which ranges from −1 to 1. The magnitude and sign of the index for a given question reflect how well that question discriminates between high- and low-scoring students; a positive value indicates that high-scoring students tended to answer the question correctly, while low-scoring students didn't, and vice-versa. A minimum acceptable correlation coefficient threshold of 0.15 is suggested, with good items generally performing at >0.25 (19). Questions performing lower than the minimal threshold should be reviewed or refined for wording, presentation, and context.

It was obvious from the examples submitted that instructors are striving for strong and innovative assessments aligned to a set of core competencies. However, it was also clear that creating effective assessments takes considerable time and effort, as NIBLSE instructors who contributed assessments often qualified their examples as "drafts" or "evolving." Here, we provide an overview of assessment types and encourage the reader to further explore the rich literature on assessment in the STEM classroom, starting with Handelsman et al. (20) and Dirks et al. (21). These 10 crowd-sourced assessment types and accompanying summary provide instructors with a quick reference for designing aligned assessment instruments independent of classroom instructional modality.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, DOCX file, 1.5 MB.

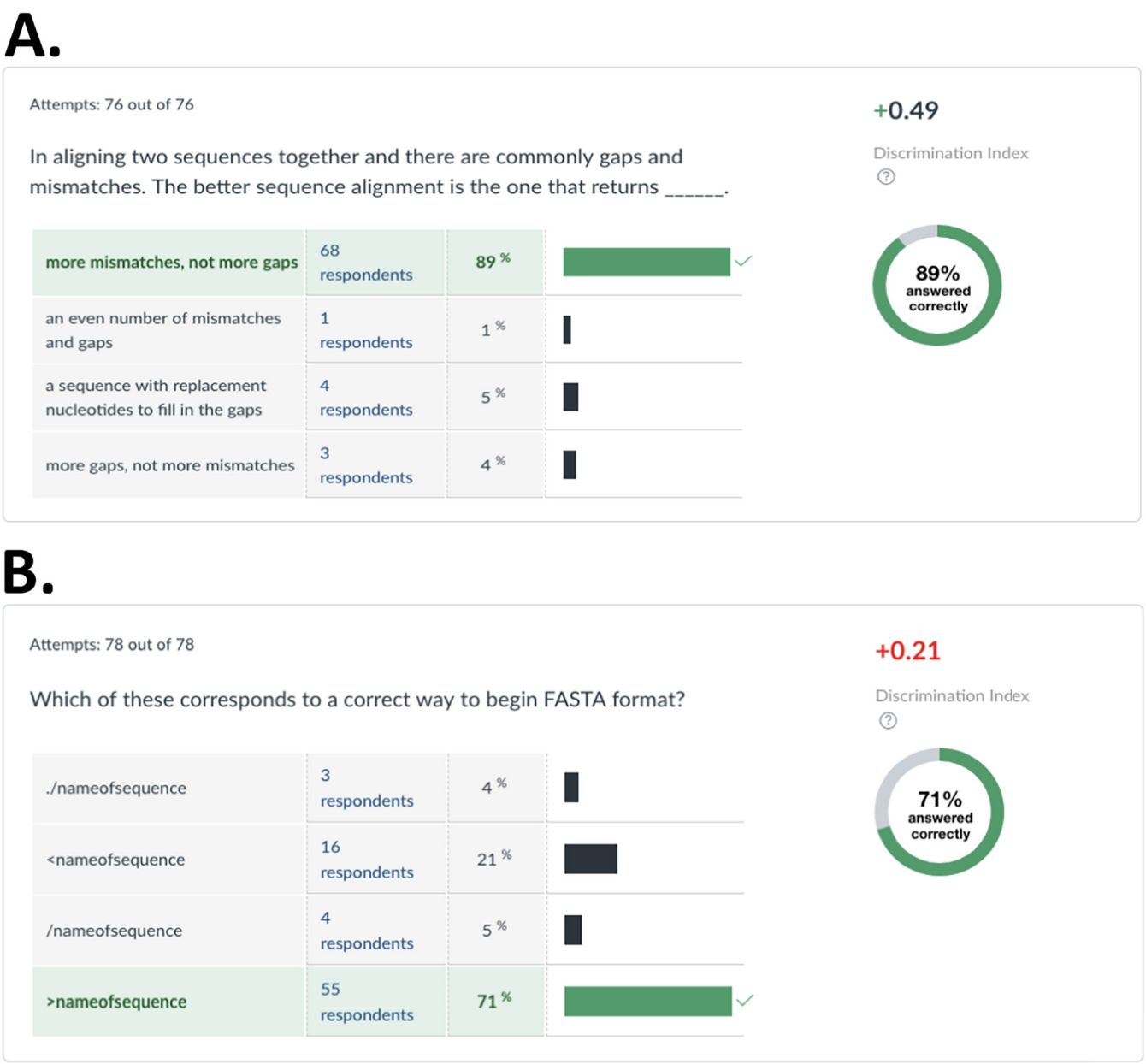

FIG 2. Screenshot of bioinformatics assessment results of two different multiple-choice questions. The question in panel A has a higher discrimination index than that in panel B, which means that it is more effective at discriminating between high- and low-scoring students. The question in panel B has a relatively low index, which suggests that the question is actually intruding upon the objective and indicates that high-performing students may be confused or "tricked" by that question. Lower discrimination indices (<0.25) are often labeled in red by the LMS to alert that the question may need to be reviewed or refined.

ACKNOWLEDGMENTS

We thank the NSF for its support of the Network for the Integration of Bioinformatics in Life Sciences Education (NIBLSE) as a Research Coordination Network for Undergraduate Biology Education (award 1539900).

We thank our colleagues who provided question sample types in the supplemental material.

We declare no financial, personal, or professional conflict of interest related to this work. The views expressed here are those of the authors and do not reflect the position of our respective organizations or supporting entities.

REFERENCES

1. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American

Association for the Advancement of Science July 15–17, 2009. Washington, DC.
2. American Association for the Advancement of Science. 2015. Vision and change in undergraduate biology education: chronicling change, inspiring the future. Washington, DC.
3. American Association for the Advancement of Science. 2018. Vision and change in undergraduate biology education: unpacking a moment and sharing lessons learned. Washington, DC.
4. Williams JJ, Drew JC, Galindo-Gonzalez S, Robic S, Dinsdale E, Morgan WR, Triplett EW, Burnette JM, Donovan SS, Fowlks ER, Goodman AL, Grandgenett NF, Goller CC, Hauser C, Jungck JR, Newman JD, Pearson WR, Ryder EF, Sierk M, Smith TM, Tosado-Acevedo R, Tapprich W, Tobin TC, Toro-Martínez A, Welch LR, Wilson MA, Ebenbach D, McWilliams M, Rosenwald AG, Pauley MA. 2019. Barriers to integration of bioinformatics into undergraduate life sciences education: a national study of US life sciences faculty uncover significant barriers to integrating bioinformatics into undergraduate instruction. PLoS One 14:e0224288. https://doi.org/10.1371/journal.pone.0224288.
5. Feldman J. 2018. Grading for equity: what it is, why it matters, and how it can transform schools and classrooms. Corwin Press, Thousand Oaks, CA.
6. Irvine V. 2020. The landscape of merging modalities. EDUCAUSE Rev. https://er.educause.edu/articles/2020/10/the-landscape-of-merging-modalities.
7. Korkmaz G, Toraman Ç. 2020. Are we ready for the post-COVID-19 educational practice? An investigation into what educators think as to online learning. IJTES 4:293–309. https://doi.org/10.46328/ijtes.v4i4.110.
8. Wilson Sayres MA, Hauser C, Sierk M, Robic S, Rosenwald AG, Smith TM, Triplett EW, Williams JJ, Dinsdale E, Morgan WR, Burnette JM, Donovan SS, Drew JC, Elgin SCR, Fowlks ER, Galindo-Gonzalez S, Goodman AL, Grandgenett NF, Goller CC, Jungck JR, Newman JD, Pearson W, Ryder EF, Tosado-Acevedo R, Tapprich W, Tobin TC, Toro-Martínez A, Welch LR, Wright R, Barone L, Ebenbach D, McWilliams M, Olney KC, Pauley MA. 2018. Bioinformatics core competencies for undergraduate life sciences education. PLoS One 13: e0196878. https://doi.org/10.1371/journal.pone.0196878.
9. Dinsdale E, Elgin SC, Grandgenett N, Morgan W, Rosenwald A, Tapprich W, Triplett EW, Pauley MA. 2015. NIBLSE: a network for integrating bioinformatics into life sciences education. CBE Life Sci Educ 14:le3. https://doi.org/10.1187/cbe.15-06-0123.
10. Quantitative Undergraduate Biology Education and Synthesis. 2021. Network for Integrating Bioinformatics into Life Sciences Education. https://qubeshub.org/community/groups/niblse. Accessed 17 February 2021.
11. Ryder EF, Morgan WR, Sierk M, Donovan SS, Robertson SD, Orndorf HC, Rosenwald AG, Triplett EW, Dinsdale E, Pauley MA, Tapprich WE. 2020. Incubators: building community networks and developing open educational resources to integrate bioinformatics into life science education. Biochem Mol Biol Educ 48:381–390. https://doi.org/10.1002/bmb.21387.
12. Carmines EG, Zeller RA. 1979. Reliability and validity assessment, vol. 17. Sage Publications, Thousand Oaks, CA.
13. Campbell CE, Nehm RH. 2013. A critical analysis of assessment quality in genomics and bioinformatics education research. LSE 12:530–541. https://doi.org/10.1187/cbe.12-06-0073.
14. Adcock R, Collier D. 2001. Measurement validity: a shared standard for qualitative and quantitative research. Am Polit Sci Rev 95:529–546. https://doi.org/10.1017/S0003055401003100.
15. Tulsky DS. 1990. An introduction to test theory. Oncology (Williston Park, NY) 4:43–48.
16. Trochim WM. 2021. Research methods knowledge base. https://socialresearchmethods.net/kb/reltypes.php. Accessed 17 February 2021.
17. Anderson LW, Krathwohl DR. 2001. A taxonomy for learning, teaching, and assessing, abridged ed. Allyn and Bacon, Boston, MA.
18. Matlock-Hetzel S. 1997. Basic concepts in item and test analysis. http://www.ericae.net/ft/tamu/Espy.htm. Accessed 1 October 2020.
19. Varma S. 2006. Preliminary item statistics using point-biserial correlation and p-values. Educational Data Systems, Inc., Morgan Hill, CA.
20. Handelsman J, Miller S, Pfund C. 2007. Scientific teaching. Macmillan, New York, NY.
21. Dirks C, Wenderoth MP, Withers M. 2014. Assessment in the college science classroom. WH Freeman, New York, NY.