How should I write exam questions? An investigation into how different framings of exam questions in biology classes can influence student performance and attitudes

Jeremy Hsu, Chapman University, Orange, CA, hsu@chapman.edu
Noelle Clark, Chapman University, Orange, CA, noclark@chapman.edu
Kate Hill, Chapman University, Orange, CAkatill@chapman.edu
Melissa Rowland-Goldsmith, Chapman University, Orange, CA, rowlandg@chapman.edu

Subject/problem: Quizzes and exams are nearly ubiquitous across both K-12 and college biology courses, with such assessments often playing major roles in determining student success and persistence in science, technology, engineering, and math (STEM). However, little work has explored how the framing of assessment questions may influence student performance and affect, despite past work showing that small changes in questions can have large impacts. For instance, personalizing questions with students' interests (i.e., grounding scenarios in students' academic and extracurricular topics relevant to students) can increase motivation and learning (e.g., Awofala 2014; Bernacki & Walkington 2014; D'Agata 2015; Ku & Sullivan 2001; Melsky 2021). However, this past work has primarily been done in the context of math, physics, and engineering courses, and we are not aware of any work examining the influence of how questions are worded in biology classes on student performance or affect.

Here, we explore question framing in scenario-based constructed-response questions where students read real scenarios and predict results in the context of an undergraduate introductory molecular genetics course. These authentic assessments mimic real-world application since students think critically about open-ended tasks (Koh 2017; Wiggins 2019). We also situate our work in discourse comprehension (Van Dijk & Kintsch 1983). Under this theory (also known as construction-integration), students must build both a textbase and situation model when reading a new scenario. The textbase represents a basic understanding of the language used and contains only minimal levels of inferences, while the situation model represents more complex mental representations (Graesser & Zwaan 1995; Gunel et al. 2009; Kintsch 1986; Van Dijk & Kintsch 1983).

Design or procedure: We had two research questions:

1. How does different framing of isomorphic assessment questions impact student comprehension, performance, and affect?
2. What framing do students prefer in scenario-based assessment questions?

We conducted a quasi-random study with three sections of an introductory molecular genetics college course with common curriculum and assessments; each section had approximately 55 students enrolled and similar demographics across the sections. However, we varied the framing of assessment questions for each section. The first section had authentic framing, using diverse scientist names ("Dr. Meiya Chen examined…") since counterstereotypical scientist names can lead to higher STEM identity and sense of belonging (Schinske et al. 2016; Sharkawy 2012; Yonas et al. 2020). The second section had self-referential framing, which uses "you" to place the reader in the experiment ("Suppose you examined…"). We chose to use this framing since students who think about themselves while encoding new information may process that information better (Conway & Dewhurst 1995; Craik & Tulving 1975; Mayer et al. 2004). The last section had classmate referential framing, where a classmate's name was used ("Veronica

examined…"). This framing was used given that past work in K-12 education suggests that using student names lowers cognitive load and increases motivation and performance, and multiple professional development sites for teachers have recommended that instructors use classmate names in assessments (Hart 1996; Riley 2001).

We measured students' STEM identity at the start of term using a pre-class survey deployed to all students and used validated Likert-scale questions in each quiz or exam to measure students' identity. In addition, we surveyed students at the end of term which variant they most and least preferred and why. After developing codebooks, we coded 30 responses and identified high interrater reliability (Cohen's kappa > 0.7 for each question). Given this high interrater reliability, one person coded the remaining responses. We also compared student affect and performance between sections and students with different variant preferences. Finally, to further validate our responses, we conducted semi-structured interviews with 11 students (4 students from two of the sections, and 3 from the remaining section) to further explore the impact of the different framings on student comprehension and affect. Interviews were transcribed and analyzed by two researchers to identify emergent themes.

Analyses and findings: We find no difference in student performance or affect across sections (ANOVAs; $p>.05$) or variant preferred (Kruskal-Wallis; $p>.05$), indicating that the different versions may not impact student performance on the assessments. However, the majority of students (62.1%) across sections indicated that they most preferred the self-referential framing. Student responses suggest that the variants impact their textbase and situation model. For instance, students cited how the self-referential variant was easier to read, facilitating textbase construction, and had less complexity, lowering cognitive load. A fourth of students indicated that it was easier to see themselves conducting the experiment with the self-referential version, forming a situation model more easily by centering a scenario around themselves (D'Ailly et al. 1995; Scheller & Sui 2022). Interviews confirmed that many students viewed the self-referential version as easier to read and more affirming of their identities when completing an assessment. Taken together, these results indicate that instructors may wish to consider using the self-referential version given the potential cognitive and affective benefits for this version if their students also prefer this version.

In contrast, the plurality of students (40.8%) indicated that they least preferred the authentic variant, a trend consistent across sections. Students gave many reasons for disliking the authentic variant. Many stated that scientist names led to increased complexity and cognitive load, hindering textbase formation. Students also cited decreased self-efficacy from seeing a scientist's name, with many intimidated by the scenario, lowering interest and motivation. Finally, students showed disparate responses for the classmate referential version. For instance, many students recognized the classmate name (or indicated that the name sounded like a peer), indicating increased confidence in answering the question after seeing the name of a peer. However, in contrast, others did not recognize the names, lowering their situational interest and impeding textbase formation.

Contribution: Our study is the first to investigate the framing of assessment questions in undergraduate biology (and the first we are aware of for any K-12 biology context as well),

providing insight into the influence on students' comprehension, performance, and affect. Our end of term survey and interviews reveal that differences in framing likely impact some students' ability to build a textbase and situation model and how they relate to the experiment. This work is of interest to both biology education researchers and practitioners. We provide new insights on the impact of assessment framing on students' affect and comprehension, informing future work on how students process questions. Our work also provides instructors insight for writing assessment questions.

General interest: Our work is of general interest to all members of NABT. For instance, instructors of primary, secondary, and post-secondary biology courses will be interested in the insights and implications for writing quiz and exam questions, which are ubiquitous across K-12 and college biology classes. Similarly, biology education researchers will be interested in the findings and ways to continue investigating the framing of assessment questions to provide additional insight into the most equitable way of framing assessment questions.

References:

Akinsola, M. K., & Awofala, A. O. A. (2009). Effect of personalization of instruction on students' achievement and self-efficacy in mathematics word problems. *International Journal of Mathematical Education in Science and Technology*, *40*(3), 389–404. https://doi.org/10.1080/00207390802643169

Awofala, A. O. A. (2014). Examining Personalisation of Instruction, Attitudes toward and Achievement in Mathematics Word Problems among Nigerian Senior Secondary School Students. *International Journal of Education in Mathematics, Science and Technology*, *2*(4), Article 4.

Bernacki, M. L., & Walkington, C. A. (2014). The Impact of a Personalization Intervention for Mathematics on Learning and Non-Cognitive Factors. *EDM (Workshops)*.

Conway, M. A., & Dewhurst, S. A. (1995). The self and recollective experience. *Applied Cognitive Psychology*, *9*(1), 1–19. https://doi.org/10.1002/acp.2350090102

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

D'Agata, B. (2015). The Influence of Personalized Mathematical Word Problems on Second Graders' Performance, Attitudes Toward Word Problems, and Difficulty Ratings. *Student Research Submissions*. https://scholar.umw.edu/student_research/118

D'Ailly, H. H., Murray, H. G., & Corkill, A. (1995). Cognitive Effects of Self-Referencing. *Contemporary Educational Psychology*, *20*(1), 88–113. https://doi.org/10.1006/ceps.1995.1005

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology*, *83*(1), 61.

Graesser, A. C., & Zwaan, R. A. (1995). Inference generation and the construction of situation models. *Discourse Comprehension: Essays in Honor of Walter Kintsch*, 117–139.

Gunel, M., Hand, B., & McDermott, M. A. (2009). Writing for different audiences: Effects on high-school students' conceptual understanding of biology. *Learning and Instruction*, *19*(4), 354–367. https://doi.org/10.1016/j.learninstruc.2008.07.001

Hart, J. M. (1996). The effect of personalized word problems. *Teaching Children Mathematics*, *2*(8), 504–506.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, *3*(2), 87–108.

Koh, K. H. (2017, February 27). *Authentic Assessment*. Oxford Research Encyclopedia of Education. https://doi.org/10.1093/acrefore/9780190264093.013.22

Ku, H.-Y., & Sullivan, H. J. (2001). *Effects of Personalized Instruction on Mathematics Word Problems in Taiwan*. https://eric.ed.gov/?id=ED470135

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, *96*(2), 389.

Melsky, K. (2021). *Effect of Personalized Problems in Undergraduate Thermal Fluid Transport Courses* [PhD Thesis]. Tufts University.

Riley, M. M. (2001). *The effects of problem revision of mathematics word problems on performance and anxiety for fourth-grade students* [Ph.D., University of South Carolina]. https://www.proquest.com/docview/250185330/abstract/59812E5E8D25481BPQ/1

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1999). *Self-reference and the encoding of personal information* (p. 149). Psychology Press.

Scheller, M., & Sui, J. (2022). The power of the self: Anchoring information processing across contexts. *Journal of Experimental Psychology: Human Perception and Performance*.

Schinske, J. N., Perkins, H., Snyder, A., & Wyer, M. (2016). Scientist Spotlight Homework Assignments Shift Students' Stereotypes of Scientists and Enhance Science Identity in a Diverse Introductory Science Class. *CBE—Life Sciences Education*, *15*(3), ar47. https://doi.org/10.1187/cbe.16-01-0002

Sharkawy, A. (2012). Exploring the potential of using stories about diverse scientists and reflective activities to enrich primary students' images of scientists and scientific work. *Cultural Studies of Science Education*, *7*(2), 307–340. https://doi.org/10.1007/s11422-012-9386-2

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Turk, D. J., Gillespie-Smith, K., Krigolson, O. E., Havard, C., Conway, M. A., & Cunningham, S. J. (2015). Selfish learning: The impact of self-referential encoding on children's literacy attainment. *Learning and Instruction*, *40*, 54–60. https://doi.org/10.1016/j.learninstruc.2015.08.001

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*.

Wiggins, G. (2019). The Case for Authentic Assessment. *Practical Assessment, Research, and Evaluation*, *2*(1). https://doi.org/10.7275/ffb1-mm19

Yonas, A., Sleeth, M., & Cotner, S. (2020). In a "Scientist Spotlight" Intervention, Diverse Student Identities Matter. *Journal of Microbiology & Biology Education*, *21*(1), 25. https://doi.org/10.1128/jmbe.v21i1.2013