

# **Supplemental Material**

*CBE—Life Sciences Education*

Hilton *et al.*

## Supplementary Material

We now present some of the comments that students gave each other during steps 4 and 5 to help show educators the kinds of discussion that using PeerWise for collaborative question authoring can elicit.

The authors AH and GD developed a coding scheme and rated all the comments that students generated. These comments were geared towards improving the question, and so not every kind of comment was useful for our current research purposes. For example, one of the codes to classify comments concerned word choice or sentence structure. Another code noted when a comment was just generally encouraging or complimentary. Although these comments were quite frequent, more substantive comments were also included, such as a code for “explain” which offered an explanation of some kind, and “confess” wherein a commenter confessed that this question made them realize they had misunderstood something. Any comment could also be coded as belonging to multiple categories, such as “confess” + “explain.” An advantage of using a large cohort of students is that even with the high number of comments that do not reveal much about student conceptual understanding, there were still hundreds of these highly substantive comments. Here is an illustrative example.

“A patient has been experiencing severe weight loss. Following a number of tests, they have been diagnosed with Type 1 diabetes. Which of the following findings most directly links to the cause of the weight loss?”

- A. Low levels of GLUT-4 translocation to plasma membrane.
- B. High levels of lactate in blood
- C. Low levels of active PFK-2
- D. High levels of cAMP in fat cells”

The explanation offered by the student who wrote the question:

“Massive weight loss in Type 1 Diabetes is primarily caused by uncontrolled lipolysis. cAMP is a stimulator of lipolysis and in a person with normal insulin levels cAMP is inhibited by insulin. Without the production of insulin cAMP is freely allowed to stimulate lipolysis in fat cells and so if this was found in a test on a diabetic person it would be the best indicator of the underlying cause of the weight loss. C is just plain wrong - insulin stimulates the activation of PFK-2 which in turn will stimulate gluconeogenesis. A and B are both also indicators of low insulin but do not directly connect to the weight loss symptom - a lack of GLUT-4 translocation explains high blood sugar content and high levels of lactate in the blood are indicative of acidosis, neither of which contribute to the weight loss in as direct a manner as the cAMP levels.”

Here is a comment about this question that was coded as a “confess.”

“Great question, I accidentally got the question wrong because I got a bit confused with whether insulin inhibited or activated the cAMP in fat cells. Now I won't make that mistake again and read more carefully! I really do like this question and believe it is a D/HD<sup>1</sup> MCQ due to its complexity of not only knowing the big pathway of the symptom but specifically requiring the knowledge of it in a deeper molecular level.”

Here are two more “confess” comments. The first is in response to a question concerning starvation’s effect on the liver:

“This question is quite difficult but your explanation justified the correct answer quite well. I made the mistake of immediately jumping to protein breakdown even if it is quite right that it is the last resort and thus the production of ketone bodies by the liver would come before it reached that point. Furthermore, the question is quite well written as it simplifies the overall response of the body after 24 hours of starvation.”

The second is in response to a question about an electron transport chain:

“This is a great question. I love how you gave the question some context; this made it a little more interesting than another bland and poorly introduced question on the electron transport chain. I also like how you, against so many others have actually provided the correct answer, as yes, there is no proton pumping occurring at complex II. You won't believe how many people still think that protons are pumped from complex II. I would not change this question. Thank you for your submission”

These “confess” comments all pertain to a misunderstanding of a molecular causal process (e.g., whether something activates vs. inhibits, or when in a sequence a causal step occurs, etc.), which as discussed in the introduction, is critical to shift from novice to expert thinking (Galloway et al., 2018). Identifying which causal processes are frequently misunderstood in the student cohort is a crucial step in designing instruction to foster expert-like causal process-based concepts. PeerWise creates a platform where these kinds of misunderstandings can be readily identified.

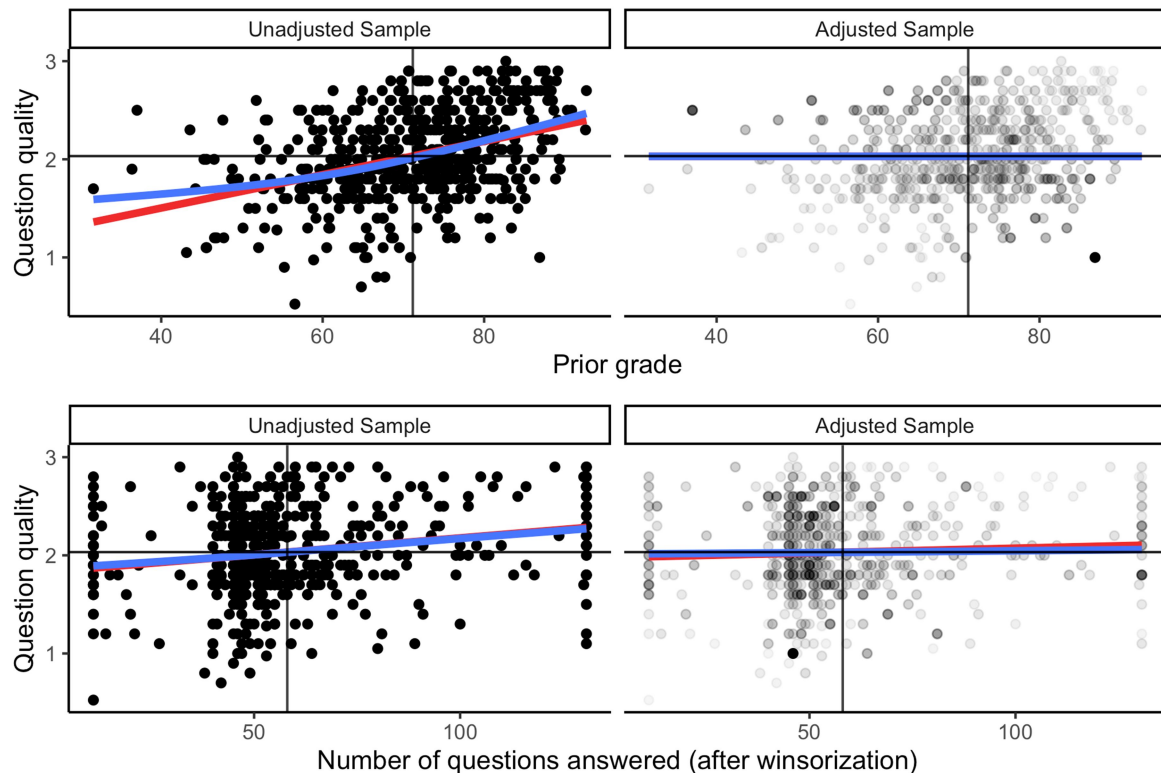
We highlight these examples so educators can get a sense that this kind of question authoring activity can engage critical concept learning processes.

---

<sup>1</sup> Within the Australian higher-education system, D = Distinction (a percentile mark of 75-85%) HD = High-Distinction (85-100%).

## Supplementary Figures

### Distributional balance across control variables



**Supplementary Figure 1:** These plots show the distributional balance for prior grades (top) and number of questions used to self-test (bottom), both with respect to the question authoring score. The plots on the left show the unadjusted values (each point corresponds to a student and their exam mark) and then after weights have been estimated using inverse-probability weighting, the final weights are visualized as the transparency of the points. Regression lines show the extent to which this procedure corrects for distributional balance (a horizontal line indicates perfect balance). The clustering of the points at the minimum and maximum of the bottom plot are a result of the winsorization approach that attenuates the influence of extreme outliers.