Why Knowledge Analysis Changes the Design of Computational Learning Environments in Biology Education

Matthew Lira, University of Iowa, mlira@uiowa.edu

Abstract: I present a pilot investigation that leveraged a Knowledge-in-Pieces approach to add theoretical specificity to characterizing the challenges and opportunities undergraduate biology students experience when coordinating knowledge of mathematical and physical models in the domain of physiology. Leveraging prior studies, I re-designed a computational modeling environment in NetLogo along with an instructional protocol to support students' capacity to readout and distinguish between two different physical quantities responsible for a dynamic equilibrium process. A between-subjects experiment offers emerging evidence that guiding students' visual attention and instructing them explicitly on how to readout physical quantities supports them in both mapping variables during quantitative problem solving and explaining physical relationships between equations.

Mathematical equations in biology

A few years ago, professional biologists experienced a quandary. Theoretical biology was failing to communicate with experimental biology (Fawcett & Higginson, 2012). Increased density of equations in theoretical papers correlated negatively with citation rates in experimental papers. The report spurred a flurry of response letters and although scholars disagreed about the takeaway, they agreed that readers needed support to interpret equations.

This event offers a lesson for designing undergraduate biology education learning environments. Communicating with mathematical representations in science demands that we understand how people coordinate representations with knowledge of a domain. Despite repeated calls for increased mathematical approaches in biology education (Bialek & Botstein, 2004), extant efforts often tap only the pedagogical and content expertise of instructors to create curricular and institutional change.

These efforts could be complimented by theory-driven approaches that model and assess *students* learning mechanisms. Such approaches would inform our understanding of the challenges and opportunities students experience and specify principles of design at the math-bio interface. The current drive to improve biology students' conceptual understanding *and* their math skill affords the opportunity to leverage theories and methods from the Learning Sciences (LS) to address this historically widespread science education concern.

Math skill learning remains divorced from conceptual learning models

Biology educators concern themselves with students' conceptual understanding. But for some, students' intuitive knowledge of biology—e.g., anthropomorphizing—conflicts with mechanistic biology (Coley & Tanner, 2015). Within this problem space, what role do math procedures play? For instance, how does computing a rate of diffusion teach a student to refrain from saying molecules "want" to diffuse across a cell membrane? Likewise, studies that assess students' math skill do not specify how these skills will change the conceptual understanding students bring to class (Speth et al., 2010). Thus, we need a model of knowledge that specifies how coordinating students' mathematical perceptions and actions might lead them to construct different conceptual understandings.

Math representations shape the conceptual understanding students construct

The division between conceptual change and math skill learning matters because emerging models of knowledge in biology do not centralize representations (Coley & Tanner, 2015). But any mathematical approach will be embodied in a semiotic system—it will express ideas in a set of representations that afford and constrain ideas and actions. Biology students are increasingly asked to interpret various mathematical representations to learn (Speth et al., 2010). When physics students learn with equations, they construct knowledge of *balance* and *equilibrium* but when they learn in computational modeling environments, they construct knowledge of *process* and *causation* (Sherin, 2001). This means that different representational systems lead students to construct different conceptual understandings. In more proximate terms, the features of representations cue different heuristics from learners (Elby, 2000). Students often do not "see" the mathematical concept or physical quantity that instructors want them to see expressed in representations. But if we teach students how to *readout* the target knowledge from the representations designed to express it, we may support them in coordinating knowledge of mathematical concepts and physical quantities in representational learning environments and support their conceptual understanding of biology.

Coordination-class theory offers BER a model of representational knowledge

Knowledge Analysis refers to a set of *methodological* approaches based upon the *theory* of Knowledge-in-Pieces (KiP). The theory generates different models of knowledge as complex systems of elements. One model, coordination-classes (Parnafes, 2007), specifies ways that students' intuitions interact with features of scientific representations. The model posits that students learn about mathematical concepts and physical quantities by coordinating *extraction and readout strategies* (i.e. how they perceptually notice features and determine information in representations) to construct inferences vis-à-vis representations. By coupling perceptual processes to conceptual understanding, the model offers new avenues to observe and direct student learning. Instead of targeting either verbal and diagrammatic instruction or else mathematical procedures with equations, a third path reveals itself. Learning how to "see" mathematical and physical quantities is increasingly recognized throughout cognitive science as a key feature of developing expertise. Delivering instruction of this nature yields positive results (Kellman, Massey, & Son, 2010). Therefore, I pose the following guiding research question:

- Does teaching readout strategies to students improve their determining mathematical and physical quantities as evidenced by problem solving and explaining physical relations between equations?

Method

Population. 7 students sampled from an advanced animal physiology course that applied quantitative approaches to learn about physical and chemical properties of physiological mechanisms.

    Disciplinary context. The resting membrane potential belongs to a class of phenomena related to dynamic equilibrium (see Figure 1). To explain the target phenomenon (i.e. a steady-state voltage), students must coordinate two factors. A *chemical concentration gradient* for an ionic species and a *selectively permeable membrane* will result in diffusion of that species along its concentration gradient and thus, will result in the separation of oppositely charged particles at the membrane (i.e. an electrical gradient). The system may be modeled ideally as when one species dominates—thus, the Nernst equation—$\frac{60}{z} \times log_{10}\left(\frac{ion_{out}}{ion_{in}}\right) = E_{ion}$—describes these conditions when the cell's membrane potential is the *same* as the equilibrium potential for a given ionic species. The right variable (E_{ion}) represents the predicted voltage that will precisely balance the left variables and parameters that represent the concentration gradient for an ionic species.

    Procedure. A between-subjects design was employed. A negative control served as a comparison. The experimental group (Readout condition) completed a non-timed pre-assessment that required calculation with problems and written descriptions of the relations between the Nernst (ideal conditions) and a related equation (non-ideal). They then received instruction with a NetLogo model (Wilensky, 1999). The instruction explained how to determine a concentration gradient and how to distinguish it from the electrical gradient with talk and using the cursor to point to fluid compartments and monitors (see Figure 1). Students then completed items similar to the pre-assessment to ensure they learned—any errors were corrected with an explicit explanation. They then completed a quantitative problem-solving test. Students in the Control group completed the test first and then experienced the same procedure as the Readout group.
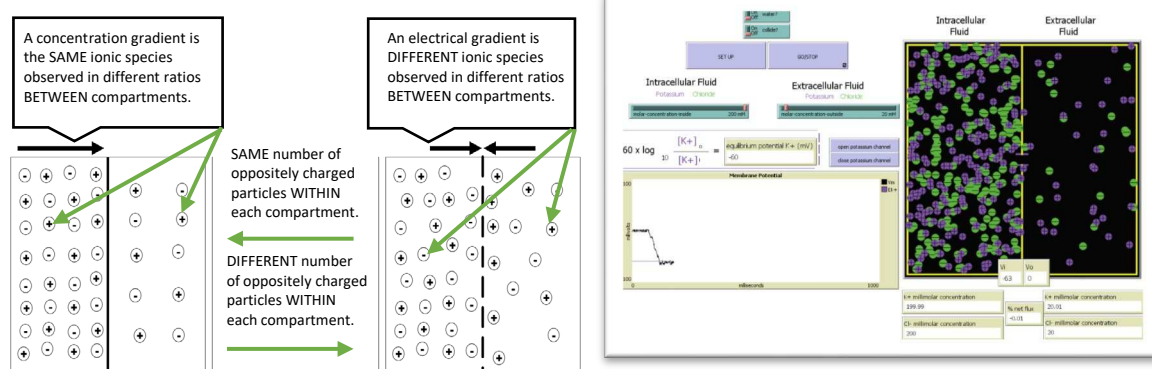

Figure 1. Left: Schematic of readout strategy instructions for distinguishing between the two physical quantities (concentration and electrical gradients). Instructions were delivered before the run (left portion) and after (right portion). Right: Screenshot of actual NetLogo simulation used during instruction.

    Instrument. The calculation items were adapted from a prior study to construct a 32-item instrument. 16 items were low complexity and 16 were high complexity (see Figure 2).
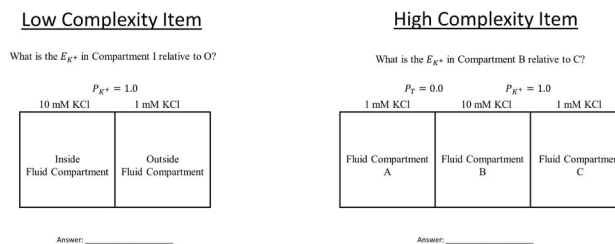
Figure 2. Sample items for low (left) and high (right) complexity.

Analysis. The study is ongoing. The present low power does not warrant inferential statistics. On the instrument, means and standard deviations were calculated to assess the emerging quantitative trends. A constant comparative analysis was used to assess students' written explanations for when and why they would use either the ideal or non-ideal equations *under ideal conditions*. 3 categories emerged (Table 1).

Results and discussion

On low-complexity items, students who did not receive readout instruction ($M=.83$ $SD=.19$) appear to perform as well as the Readout group ($M=.84$ $SD=.19$). On high-complexity items, however, students who received the Readout instruction ($M=.86$ $SD=.18$), appear to be gaining a protective effect compared to students who did not receive readout instruction ($M=.75$ $SD=.23$). This preliminary pattern supports the Readout Hypothesis (see Figure 3). Learning how to distinguish between different physical quantities appears to modify students' procedural calculation skill. Does it support their conceptual understanding of how equations model a system?
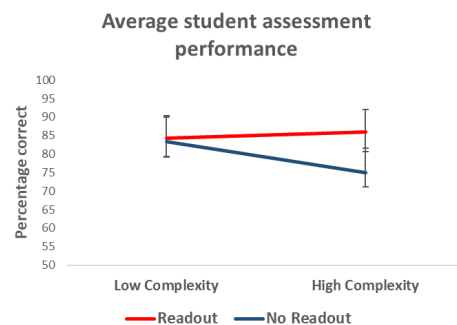

Figure 3. Student performance is trending towards Readout protecting against problem solving complexity.

Yes, it seems. Before and after readout instruction I directly assessed students' skill at explaining the relations between equations. Students shifted from describing calculation procedures (e.g., "I would use the Nernst equation to *first calculate E_{K^+}* […]) to offering accurate and meaningful physical relations (e.g., The membrane potential would be *the same as* the Nernst potential […]) that specified *mathematical equivalency under ideal conditions* (See Table 2). This shift indicates that students are beginning to see correctly that different mathematical symbols in equations can represent the *same* physical quantity depending upon how we model the system. This skill reflects a desired instructional target for modeling biological systems with math.

Table 1: Students shifted away from describing calculations and towards coordinating the equations' meanings

Written response score frequencies		
Level	Before simulation	After simulation
Calculation description	7	7
Quantitative dependency	7	2
Mathematical equivalency	0	5

In sum, prior work illustrated that students used a heuristic akin to *more x means more y* as observed elsewhere (Elby, 2000). Students saw more concentration gradient as meaning more electrical gradient despite these reflecting distinct physical quantities. This study promises that learning to coordinate knowledge of mathematical and physical models by learning to readout and distinguish between two different extraction strategies offers a viable instructional strategy. This study is being further powered to improve its rigor and generalizability.

Conclusions and implications: Towards computational biology education

Mathematical approaches in biology education will necessitate new modes of learning as students must construct new knowledge vis-à-vis new *representations*. Guiding students to learn *with* representations has been an area of interest central to LS (Elby, 2000). Teaching students how to readout physical quantities from multi-representational simulations can support their knowledge coordination as evidenced by their descriptions of mathematical relationships between equations *and* their quantitative problem solving. Thus, instruction that organizes biology students' *perception* offers an additional path for them to construct an understanding of physical quantities (Parnafes, 2007). If true, we may eschew procedural-conceptual divisions and synthesize an integrative model of representational learning for biology education (cf. Sherin, 2001).

If developing readout strategies does support conceptual understanding in quantitative biology, then what does that mean for the future of undergraduate biology education? Biology students do not typically use methods of proof. Instead, educators expect them to learn *with* computational environments like NetLogo. But why not have the students *create* the models? The approach has merit—it supports mechanistic reasoning in high school students (Wilensky & Reisman, 2006). One reason is that readout strategies become part of a hermeneutic circle—students build, run, and interpret their simulations iteratively. This offers a viable pathway for students to construct a blend of mechanistic reasoning and computational skill (Sherin, 2001). Moreover, mathematical models hold assumptions (e.g. the idealization of the Nernst equation). As opposed to, say, computing with large data sets where experimental assumptions may be unknown, students could specify their assumptions in their code. Though possible, undergraduate biology education does not yet appear to have the capacity to scale such endeavors but a growing community strives towards such a vision. In the interim, we may for now conclude that Knowledge Analysis motivates changes in the design of *mathematical* learning environments in undergraduate biology education because it focuses our attention away from a procedural-conceptual division and toward perceptual-heuristic coupling that is cued and mediated by *computational* representations.

References

Bialek, W., & Botstein, D. (2004). Introductory Science and Mathematics Education for 21st-Century Biologists. *Science*, *303*(5659), 788–790. https://doi.org/10.1126/science.1095480

Coley, J. D., & Tanner, K. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBE Life Sciences Education*, *14*(1), 1–19. https://doi.org/10.1187/cbe.14-06-0094

Elby, A. (2000). What students' learning of representations tells us about constructivism. *Journal of Mathematical Behavior*, *19*(4), 481–502. https://doi.org/10.1016/S0732-3123(01)00054-2

Fawcett, T. W., & Higginson, a. D. (2012). Heavy use of equations impedes communication among biologists. *Proceedings of the National Academy of Sciences*, *109*(29), 11735–11739. https://doi.org/10.1073/pnas.1205259109

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, *2*(2), 285–305. https://doi.org/10.1111/j.1756-8765.2009.01053.x

Parnafes, O. (2007). What does "Fast" mean? Understanding the physical world through computational representations. *Journal of the Learning Sciences*, *16*(3), 415–450. https://doi.org/10.1080/10508400701413443

Sherin, B. L. (2001). A comparison of programming languages and algebraic notation as expressive languages for physics. *International Journal of Computers for Mathematical Learning*, *6*(1), 1–61. https://doi.org/10.1023/A:1011434026437

Speth, E. B., Momsen, J. L., Moyerbrailean, G. A., Ebert-May, D., Long, T. M., Wyse, S., & Linton, D. (2010). 1, 2, 3, 4: Infusing quantitative literacy into introductory biology. *Cell Biology Education*, *9*, 323–332. https://doi.org/10.1187/cbe.10

Wilensky, U. (1999). *NetLogo*. Retrieved from http://ccl.northwestern.edu/netlogo/

Wilensky, U., & Reisman, K. (2006). Thinking Like a Wolf, a Sheep, or a Firefly: Learning Biology Through Constructing and Testing Computational Theories—An Embodied Modeling Approach. *Cognition and Instruction*, *24*(2), 171–209. https://doi.org/10.1207/s1532690xci2402_1