

MODELING SYSTEMS IN BIOLOGY: EXPLORING STUDENT CONCEPTUALIZATION OF GENE EXPRESSION AND NATURAL SELECTION

SUBJECT/PROBLEM

Vision and Change (Brewer and Smith, 2011) and the NRC Framework for K-12 Science Education (NRC, 2012) identified systems and models as a core concept and a cross-cutting competency, respectively, for biology education. Systems and models, however, have received relatively less attention in the biology education literature, compared to other concepts and practices, particularly at the college level (Treibergs et al., personal communication). Our research seeks to develop tools that support systems-focused instructional practices in undergraduate biology, while also using student-generated models to explore students' development of systems thinking skills.

Model-building helps learners make connections between concepts, and promotes causal and mechanistic reasoning (Louca and Zacharia, 2012). We use model-based pedagogical approaches to promote learning in the undergraduate introductory biology classroom, leveraging models and modeling as powerful tools for helping students reason and learn about complex biological processes (Wilson et al., 2020). Our approach is grounded in the Biological Systems Thinking (BST) framework (Momsen et al., 2022), which provides a systems-centered conceptual framing for curriculum organization, and connects systems thinking skills to classroom practices like model-building and explanation. In our courses, we teach students to represent their understanding of complex biological systems by drawing conceptual models based on the structure-behavior-function theoretical framework (Hmelo-Silver et al., 2007, 2015; Vattam et al., 2011). Briefly, this framework posits that systems have functions, resulting from complex networks of interactions (behaviors) among the systems' physical components (structures). Students learn to identify the structures and interactions or relationships that are relevant to explaining how systems function, and create simplified box-and-arrow models where structures are nouns in boxes, connected by labeled arrows indicating their relationships (Figure 1; Dauer et al., 2013; Bray Speth et al., 2014).

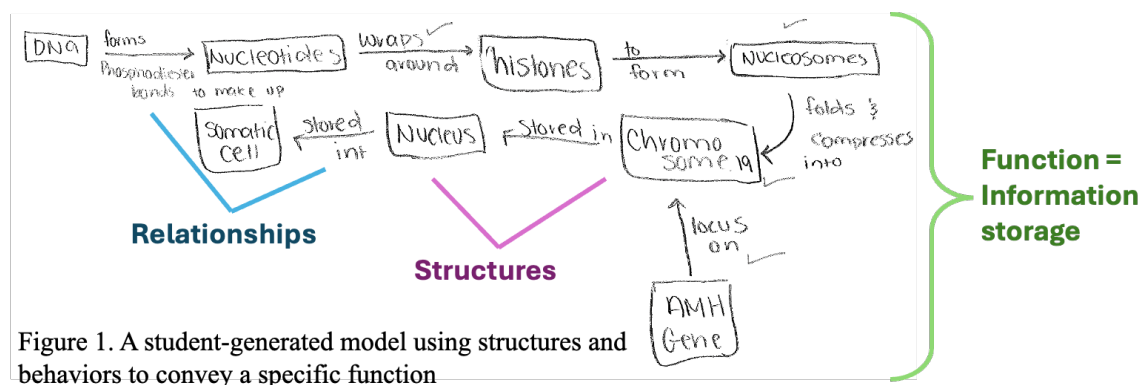


Figure 1. A student-generated model using structures and behaviors to convey a specific function

Students generate concept models as part of class activities, homework, and on exams. A key component of model-based pedagogy is iteration: students' models start small, generally with 3-5 structures and a basic function (e.g., how DNA stores information), and grow over the course of the semester, as students add structures and functions, but also revise model organization and wording based on self-evaluation, and on peer and instructor feedback.

We present two research studies, conducted at different institutions, as examples of ways in which modeling can support students' development of systems thinking skills while also revealing student thinking about two foundational biological processes often taught separately: gene expression in the cell (the central dogma of molecular genetics), and natural selection.

MODELING SYSTEMS IN BIOLOGY: EXPLORING STUDENT CONCEPTUALIZATION OF GENE EXPRESSION AND NATURAL SELECTION

Study 1 is guided by cognitive schema theory (Rumelhart & Norman, 1976), which posits that learners gradually construct knowledge through the active cognitive processes of accretion, tuning, and restructuring. We explored how introductory biology students' iterative models of gene expression reveal their knowledge development in the short time of a course unit.

Study 2 explores students' knowledge retention of the conceptual connections between gene expression and natural selection, from introductory biology to the beginning of an upper-division course on evolution. We qualitatively examined student-generated models to explore what structures and relationships remained constant or changed across time.

DESIGN OR PROCEDURE

Study 1. We examined short-term cognitive dynamics within students' conceptual models of gene expression in two sections of an introductory biology course (n=209). As they learned about the central dogma of molecular genetics, students constructed three consecutive models (over the course of 2 weeks) representing how information stored in a gene is used in the cell to produce a protein: a first-draft homework model (M1); a revised homework model after class discussion and instructor clarifications (M2); and an isomorphic model on the subsequent course exam (M3). We transcribed all models to spreadsheets as lists of *propositions* (where a proposition is the smallest meaningful segment within a model, a pair of boxes linked by an labeled arrow indicating the relationship between two structures). We analyzed student models in terms of the total number and identity of propositions, and compared successive iterations to identify what kinds of actions students took when revising their models. Each bout of revision (M1 to M2, and M2 to M3) was thus characterized by a string of values representing the student's initial model size (as total number of propositions) and the number of propositions they conserved (C), removed (R), added (A) or modified (M) in their subsequent model.

Study 2. We used a longitudinal approach to explore the impact of model-building on students' retention of key concepts from Introductory Biology to a 300-level Evolution course. Introductory Biology II serves students in their first or second year of college, and focuses on the function of information storage and flow, including the central dogma of biology and natural selection. Course instructors across all three course sections (Spring 2024) used an iterative, model-based pedagogy where students were taught to create concept models of living systems and iterated on these models across the semester. Approximately a year later, we assessed students enrolled in a 300-level course on evolution prior to instruction on natural selection. Our data streams include: (1) from Intro Bio II, student-generated models (n=266) of gene expression, contextualized to Kisspeptin and hypogonadism; (2) from Evolution, student-generated models (n=22) of gene expression contextualized to eumelanin production in rock pocket mice. Of the 22 students who completed the assessment in Evolution, eight students overlapped with our S24 introductory course. Student models were digitally transcribed using CmapTools; propositions (n=2300 for Intro Bio; n=230 for Evolution) were coded for correctness where 1 was incorrect and 3 was completely correct. Due to the small sample size (8 students), we qualitatively analyzed student models focusing on key concepts included and excluded, and accuracy of connections.

ANALYSES AND FINDINGS

MODELING SYSTEMS IN BIOLOGY: EXPLORING STUDENT CONCEPTUALIZATION OF GENE EXPRESSION AND NATURAL SELECTION

Study 1: Initial findings reveal that student models of gene expression, in aggregate, underwent changes that may be representative of initial growth by accretion, accompanied and followed by tuning and restructuring, which included significant trimming leading models to become more parsimonious over time. We applied a network visualization approach to generate meta-models of M1, M2, and M3 as whole-class aggregates (example in Figure 2): these visualizations offer a clear representation of the knowledge restructuring and simplifying process characterizing student models' change over time.

Hierarchical cluster analysis, applied to individual students' model changes, identified distinct patterns of model revision among learners. Cross-tabulation of model-revision patterns from the first bout of revision (from M1 to M2) and the second (from M2 to M3) revealed a wide range of cognitive pathways individual learners took while learning about gene expression.

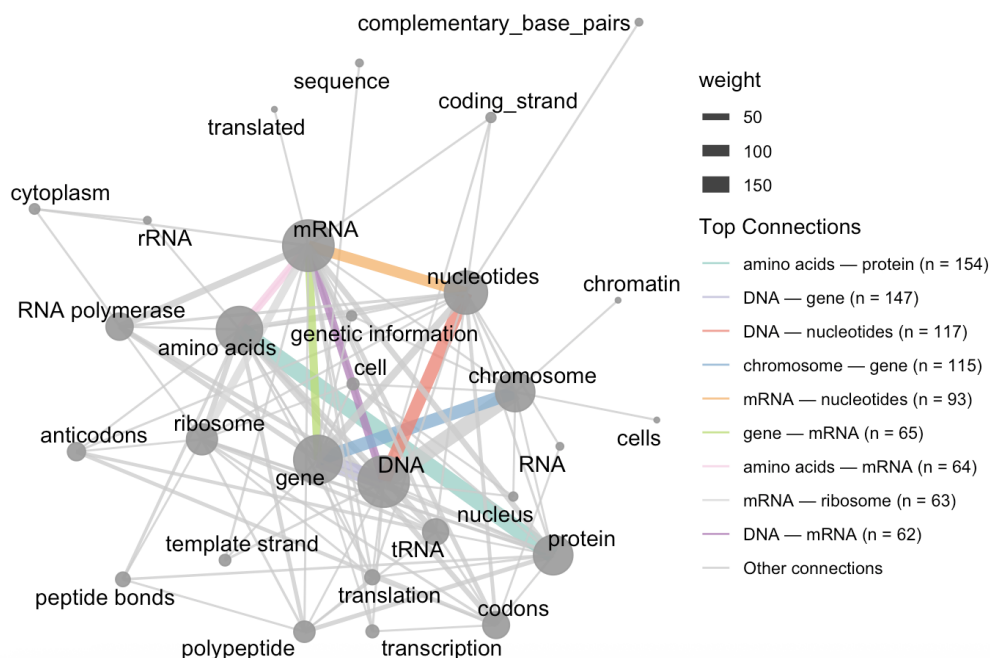


Figure 2. Example of a network representation for M3, combining all students' models (n=209); bubble sizes are proportional to the frequency of structures, and the thickness of connecting lines is proportional to the frequency of relationships (also reported in the legend for the 10 most frequent relationships).

Study 2. Students in Intro Bio created models with an average of 16 propositions, while Evolution models averaged 10 propositions; proposition correctness was also higher in the Intro course (Figure 3). Over time, student models demonstrate substantial loss of conceptual connections, possibly reflecting forgetting, but also providing evidence that students are engaged in pruning (Dauer et al., 2013), the process of dropping extraneous structures and relationships. Of the eight students we tracked from Intro Bio to Evolution, seven accurately connected the central dogma to natural selection in their Intro models. However, only one student maintained this accurate connection in Evolution, one year later. Additional analysis revealed a widespread misplacement of mutation within models, with students incorrectly locating mutation as occurring within transcription or translation processes (consistent with Bray Speth et al., 2014).

MODELING SYSTEMS IN BIOLOGY: EXPLORING STUDENT CONCEPTUALIZATION OF GENE EXPRESSION AND NATURAL SELECTION

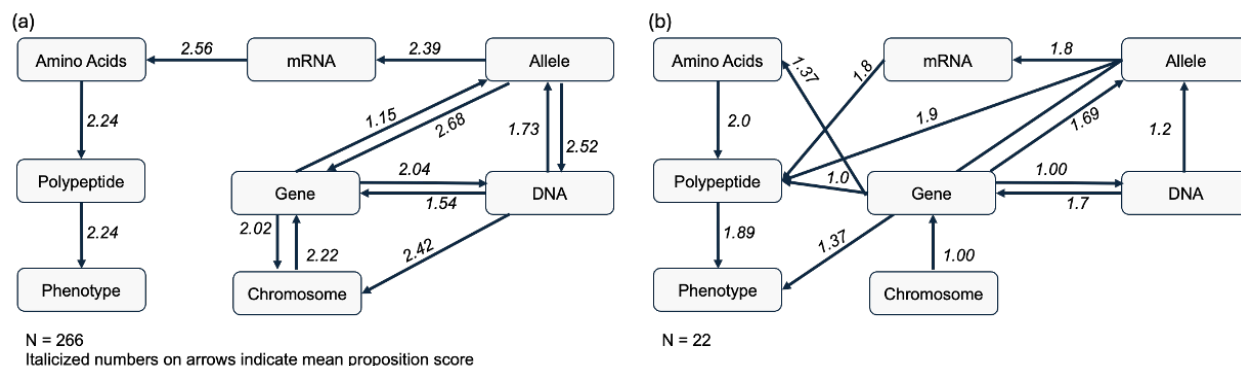


Figure 3. Mean proposition correctness (on a scale of 1-3) for models in Intro Bio II (a) and Evolution (b).

CONTRIBUTION

This research begins to fill the gap in biology education literature by exploring how modeling supports students' learning of biology and development of systems thinking.

First, through a fine-grained analysis of how students' conceptual models of gene expression evolve through iterative model-building activities, we begin to identify patterns in how students develop systems thinking skills, such as identifying the structures and interactions that are relevant to a system's function, and organizing these components to explain a biological function (e.g., protein synthesis) or process (e.g., natural selection).

Second, we document substantial conceptual forgetting and disconnection between the central dogma and natural selection over time, challenging assumptions about knowledge retention between courses. However, our findings also reveal nuance in what students remember: so-called big ideas remain, while details are lost.

Finally, this research informs curricular design. We note that traditional curricular structure separates related concepts in biology, which may impede students' ability to develop integrated knowledge structures of biological systems. The BST framework offers a path to conceiving a systems-centered curriculum, intentionally connecting biological concepts typically taught in isolation. Modeling supports this approach by offering students and instructors a transparent process to both visualize internalized schemas and connect seemingly disparate concepts.

GENERAL INTEREST

The integration of model-based learning with systems thinking aligns with current national frameworks for biology education (Brewer and Smith 2011, NRC 2012). This work contributes to collective pedagogical knowledge about how to incorporate systems and models in the biology classroom. Our analysis of student revisions illustrates a wide range of individual differences in the learning process, as students construct their knowledge about complex ideas they may be integrating for the first time. Additionally, the documentation of knowledge retention issues between courses highlights the need for greater vertical alignment within biology curricula. The findings have immediate implications for course design and concept integration, within courses and vertically, across the curriculum.

REFERENCES CITED

MODELING SYSTEMS IN BIOLOGY: EXPLORING STUDENT CONCEPTUALIZATION OF GENE EXPRESSION AND NATURAL SELECTION

- Bray Speth, E., Shaw, N., Momsen, J., Reinagel, A., Le, P., Taqieddin, R., & Long, T. (2014). Introductory biology students' conceptual models and explanations of the origin of variation. *CBE—Life Sciences Education*, 13(3), 529-539.
- Brewer, C. A., & Smith, D. (2011). Vision and change in undergraduate biology education: a call to action. *American Association for the Advancement of Science, Washington, DC*, 81.
- Dauer, J. T., Momsen, J. L., Bray Speth, E., Makohon-Moore, S. C., & Long, T. M. (2013). Analyzing change in students' gene-to-evolution models in college-level introductory biology. *Journal of Research in Science Teaching*, 50(6), 639-659.
- Hmelo-Silver, C., Liu, L., Gray, S., & Jordan, R. (2015). Using representational tools to learn about complex systems: A tale of two classrooms. *Journal of Research in Science Teaching*, 52(1), 6–35.
- Hmelo-Silver, C., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences*, 16(3), 307–331.
- Louca, L. T., & Zacharia, Z. C. (2012). Modeling-based learning in science education: Cognitive, metacognitive, social, material and epistemological contributions. *Educational review*, 64(4), 471-492.
- Momsen, J., Bray Speth, E., Wyse, S., & Long, T. (2022). Using systems and systems thinking to unify biology education. *CBE—Life Sciences Education*, 21(2), es3.
- National Research Council, Division of Behavioral, Social Sciences, Board on Science Education, & Committee on a Conceptual Framework for New K-12 Science Education Standards. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.
- Rumelhart, D. E., & Norman, D. A. (1976). *Accretion, tuning and restructuring: Three modes of learning* (No. 7602).
- Vattam, S. S., Goel, A. K., Rugaber, S., Hmelo-Silver, C. E., Jordan, R., Gray, S., & Sinha, S. (2011). Understanding complex natural systems by articulating structure-behavior-function models. *Journal of Educational Technology & Society*, 14(1), 66-81.
- Wilson, K. J., Long, T. M., Momsen, J. L., & Bray Speth, E. (2020). Modeling in the classroom: Making relationships and systems visible. *CBE—Life Sciences Education*, 19(1), fe1.