International Journal of Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Revision as an essential step in modeling to support predicting, observing, and explaining cellular respiration system dynamics

Lyrica Lucas, Tomáš Helikar & Joseph Dauer

To cite this article: Lyrica Lucas, Tomáš Helikar & Joseph Dauer (2022): Revision as an essential step in modeling to support predicting, observing, and explaining cellular respiration system dynamics, International Journal of Science Education, DOI: 10.1080/09500693.2022.2114815

To link to this article: https://doi.org/10.1080/09500693.2022.2114815

Published online: 28 Aug 2022.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
https://doi.org/10.1080/09500693.2022.2114815

Revision as an essential step in modeling to support predicting, observing, and explaining cellular respiration system dynamics

Lyrica Lucas [a], Tomáš Helikar [b] and Joseph Dauer [a]

[a]School of Natural Resources, University of Nebraska, Lincoln, NE, USA; [b]Department of Biochemistry, University of Nebraska, Lincoln, NE, USA

ABSTRACT

Comprehensive understanding of complex biological systems necessitates the use of computational models because they facilitate visualisation and interrogation of system dynamics and data-driven analysis. Computational model-based (CMB) activities have demonstrated effectiveness in improving students' understanding and their ability to use and reason with models. To maximise the effectiveness of computational modelling, this study examined an improved cognitive scaffolding and its impact on student learning of cellular respiration. This scaffolding proposes the *predict-observe-revise-explain* (PORE) sequence of tasks that explicitly challenge students to revise their predictions and computational models to resolve cognitive conflict. Based on revision work in a CMB activity, a sample of $n = 362$ undergraduate biology students were categorised into three groups – not expected to revise (NR, $n = 109$), required-revised (RR, $n = 179$), and required-did not revise (RDNR, $n = 74$). Students' performance in *predict*, *revise*, and *explain* tasks were significantly associated with post-test performance. RR students were more than twice as likely to demonstrate a positive learning gain in the post-test (*odds ratio* = 2.47) compared to RDNR students. While science education has implicitly acknowledged revision as a critical cognitive process in modelling, this study presents evidence that making revision an explicit cognitive task in a CMB activity supports student learning of a complex biological system.

ARTICLE HISTORY
Received 14 January 2022
Accepted 16 August 2022

KEYWORDS
Computational model-based activity; predict-observe-explain (POE) strategy; undergraduate life sciences education

Introduction

Computational modelling has become a central feature of scientific investigations to supplement both theoretical and experimental work as an imperative methodological approach (Bartocci & Lió, 2016; Brodland, 2015). In undergraduate life science education, computational modelling experiences are essential for understanding complex biological systems (Helikar et al., 2015; Sweeney & Sterman, 2007) because modelling is the way of knowing in science and the method by which scientific knowledge is generated (Gilbert, 1991; Lehrer & Schauble, 2000; Nersessian, 2009; Windschitl et al., 2008).

CONTACT Joseph Dauer joseph.dauer@unl.edu School of Natural Resources, University of Nebraska, Lincoln, NE, USA

Specifically, modelling can support the development of cognitive skills that enhance the acquisition of domain-specific conceptual knowledge and scientific reasoning (Buckley et al., 2004; Clement, 2000; Doerr, 1997). As computational modelling resources become increasingly available for the educational purposes of data manipulation and visualisation (de Jong & van Joolingen, 1998; Rutten et al., 2012), embedding quantitative activities into existing biology courses has transformed traditional, descriptive biology to interdisciplinary, quantitative biology curricula to address complex problems (NRC, 2003; Usher et al., 2010). This trend in life science education is expected to continue and expand since other areas of science such as physics, with its long-standing tradition of using mathematical techniques, have demonstrated the use of computer simulations as a fundamental aspect of research and education for decades (Perkins et al., 2006; Sarma & Faundez, 2017).

Biological systems are often the targets of computational modelling because learning environments that take advantage of simulations offer affordances for developing a coherent understanding of complex systems. The use of computer simulations has made investigations of biological systems on a wide scale possible and resulted in greater access to opportunities for data processing, visualisation, and testing of models (de Jong, 2006; de Jong & van Joolingen, 1998; Quintana et al., 2004). Learners benefit from the use of computer-based modelling tools in environments that prompt them to engage in the practices of generating hypotheses from partially complete models (Mulder et al., 2016), assessing tentative theories by comparing models and concurrent experimental results (Fuhrmann et al., 2018), confronting ideas with evidence and correcting errors (Wijnen et al., 2015), and collaborating and co-constructing knowledge with peers (van Joolingen et al., 2005; Wilkerson et al., 2017). There is a growing body of literature showing that these modelling practices are associated with desirable learning outcomes. Computational model-based (CMB) activities have shown that computational modelling supports undergraduate life science students' explanatory and evaluative reasoning, improves foundational systems thinking skills and conceptual knowledge, and enhances students' ability to retrieve their conceptual models (Bergan-Roller et al., 2018; Dauer & Long, 2015; King et al., 2019). Although these affordances point to the need for extensive curricular integration and increased implementation of CMB activities in life science education (AAAS, 2011; Qudrat-Ullah, 2010; Svoboda & Passmore, 2013), the use of simulations to study biological complexity is still in the early stages (Sarma & Faundez, 2017), which could explain why there are few computational modelling experiences for post-secondary students (Mulder et al., 2016; Seel, 2017).

The aim of this study is two-fold. First, we present the implementation of a CMB activity in cellular respiration in an undergraduate life science course to add to the increasing body of knowledge on how computational modelling facilitates the learning of complex biological systems. Second, we build upon the literature on *predict-observe-explain* (POE) learning strategy to examine the cognitive processes that are essential in guiding students to develop an increasingly sophisticated understanding of a complex biological system. We implemented a CMB activity in which students followed a *predict-observe-revise-explain* (PORE) sequence of cognitive tasks to scaffold their computational modelling experience. We find that the lack of emphasis on the cognitive task of revision in scientific investigations is exemplified in the commonly known POE learning strategy. In the following sections, we discuss how the PORE sequence of cognitive tasks can facilitate modelling-based learning and why revision was conceptualised as an essential cognitive task to support error detection in the process of modelling.

Embedding cognitive tasks in CMB learning activities

In CMB activities, sense-making is emphasised as students construct and modify models of biological systems that can be simulated with a computer (Penner, 2000). The parameters of a mathematical model are controlled and manipulated to examine different possible outcomes and study the dynamical behaviour of complex systems (NIBIB, 2018; Seel, 2017; Sins et al., 2005). One such complex and fundamental system that students encounter in undergraduate life science courses is cellular respiration (Bergan-Roller et al., 2020; Booth et al., 2021). Students struggle with understanding the multiple biochemical steps and molecular components. They often end up with a fragmented understanding of functions and specialisations of cell organelles and a vague conception of interconnections between the series of fundamental biochemical reactions in the system (Songer & Mintzes, 1994; van Mil et al., 2013). Learning occurs when specific cognitive processes for knowledge construction are activated (Krajcik et al., 2002; NRC, 2000b). Students' understanding of scientific phenomena become more sophisticated through the iterative practice of cognitive processes (Seel, 2017). These cognitive processes typically include identifying variables and relationships, formulating hypotheses or models, making predictions, changing variable values, interpreting outcomes, reaching conclusions, reflecting upon the learning process and newly acquired knowledge, planning inquiry, and monitoring developing knowledge (de Jong, 2006).

Although CMB activities present an array of potential benefits for students, the integration of software tools in the curriculum does not guarantee improved student learning (Krajcik & Mun, 2014). Providing students access to computational modelling resources makes new learning opportunities possible but students may also find cognitive tasks overwhelming due to heavy demands on working memory (Kirschner et al., 2006). Cognitive scaffolding is necessary to enhance student learning compared to free exploration that often leads to confusion and frustration (Alfieri et al., 2011; Lazonder & Harmsen, 2016). Thus, the design of a CMB activity should integrate computational tools and scaffolding for cognitive tasks to foster learning by modelling.

Adding revision to the predict-observe-explain (POE) learning strategy

To support modelling-based learning, we implemented a CMB activity in a collaborative setting in an undergraduate life science classroom. The structure of the CMB activity was motivated by the *predict-observe-explain* (POE) strategy developed by White and Gunstone (1992) to uncover individual students' initial ideas about a phenomenon and understand their reasoning. The POE strategy has demonstrated effectiveness in helping students develop scientific explanations by first making a hypothesis about a phenomenon, defending their hypothesis, making an observation by experimenting, and then explaining the phenomenon using data from their observations (Cinici & Demir, 2013; Haysom & Bowen, 2010; Liew & Treagust, 1998; Zacharia, 2005).

In a modelling-based learning approach, students assess their understanding when observing and detecting errors in their models and prior knowledge (Fugelsang & Dunbar, 2005). Evaluating initial models should not only result in further discussion but also in corresponding revisions of experiments and predicted outcomes that demonstrate a model-based understanding of a phenomenon (Louca & Zacharia, 2012; Schwarz et al., 2009). The POE strategy lacks an emphasis on error detection that should occur after testing a hypothesised model (Wijnen et al., 2015). Error detection is a primal neural process that disciplinary experts readily engage in when determining model plausibility (Nenciovici et al., 2019). Variations of the POE strategy expect students to reflect upon their work. However, this is only implied through the addition of activities that might lead students to engage in error detection. For instance, the POE strategy has been implemented in different ways using related sequences of cognitive tasks such as *predict-observe-explain-explore* (POEE) and *predict-discuss-explain-observe-discuss-explain* (PDEODE) to assist students in improving their conceptual understanding and problem-solving skills (Coştu et al., 2012; Savander-Ranne & Kolari, 2003). These modifications encourage learners to further explore and discuss their ideas after having the opportunity to observe or gather data. The added tasks to the basic POE strategy suggest a need to further guide students in assessing and communicating their learning. Thus, in the CMB activity used in the present study, *revision* was conceptualised as an essential cognitive task of POE, and we structured the modelling activity using a *predict-observe-revise-explain* (PORE) sequence of cognitive tasks.

The PORE strategy offers scaffolding for cognitive tasks that students typically find difficult to accomplish without external support. Although cognition is implicated in the model-building process (MacLeod & Nersessian, 2018), the PORE scaffolding makes the cognitive tasks explicit for students. This may contribute to students' control of their own learning as they engage in an iterative modelling process to understand and eventually explain the phenomenon being investigated (Papaevripidou & Zacharia, 2015).

Applying the PORE sequence in university biology

The CMB activity used in this study focused on investigations analysing context-rich phenomena and building scientific explanations supported by computational models. For example, students were asked to explain how the inhibition of one of the cellular respiration processes would affect the rest of the system resulting in written explanations about the causal implications of the inhibition. The PORE sequence of cognitive tasks in the CMB activity guided students to examine, critique, and modify their understanding of cellular respiration.

The purpose of *prediction* tasks was to enhance learning by making students cognizant of their initial understanding of the entities and activities in the processes of cellular respiration. Formally communicating predictions about how a computational model would behave provided students with a reference point when monitoring their changing conceptions through the modelling process (Lesh & Doerr, 2000). The *observation* tasks guided students through investigations and engaged them in the practices of graph reading and translation between quantitative interpretation and model representations (Mayes et al., 2014). Making predictions focused students' attention on relationships between cellular respiration processes and primed them to observe specific graphical and mathematical effects as a result of manipulating variables in their computational models. The *explanation* tasks served as a cognitive follow-through requiring students to apply higher-order thinking to analyse and synthesise their learning from model-based investigations.

The *predict-observe* sequence of cognitive tasks can bring about 'cognitive conflict' (Piaget, 1985), which is necessary to foster conceptual change. Higher-level reasoning and scientific expertise are associated with the ability to address cognitive conflicts (Bartley et al., 2019; Brookman-Byrne et al., 2018; Pennycook et al., 2015; Wijnen et al., 2015). The discrepancy between prior knowledge and experimental results presents a conflict between two different cognitive structures (Hewson & Hewson, 1984) that need to be reconciled through *revision*, such as repeating a computational experiment or changing expected outcomes. While cognitive conflict is necessary, alone, it is insufficient for conceptual change. The *revision* component is integral to the POE approach to prompt students to reflect on their computational models and prior knowledge when they experienced inconsistencies stimulated by conflicting predictions and evidence from the computational experiment (Fuhrmann et al., 2018; Krajcik & Merritt, 2012).

Research questions

The CMB activity on cellular respiration was used in a prior study and has demonstrated improvement in students' conceptual knowledge of cellular respiration and systems thinking skills based on pre – and post- concept models drawn by students to represent their mental models (Bergan-Roller et al., 2018). In the present study, we focused on investigating the impact of specific cognitive tasks in the CMB activity to explain how the CMB activity facilitated conceptual change.

Overall, this study investigated how undergraduate biology students' performance in cognitive tasks – *prediction*, *observation*, *revision*, and *explanation* - is associated with learning outcomes (i.e. post-test score and learning gains) in a CMB activity on cellular respiration. The overarching research aim was to determine cognitive tasks that are associated with positive learning outcomes in a CMB activity. To achieve this aim, we investigated the following questions: (1) Does revision work influence students' learning gain? and (2) To what extent do cognitive tasks in a CMB activity applied in a PORE sequence predict student post-test performance in cellular respiration after controlling for student differences?

Methods

This study was conducted in one semester in a large-enrolment introductory biology lab course run by three instructors in a public, research-intensive Midwestern USA university. Students who participated in the study came from 41 lab sections, each guided by a teaching assistant (usually a graduate student). The modelling-based learning activity on cellular respiration used in this study is available in Cell Collective (https://cellcollective.org; Helikar et al., 2012, 2015), a web-based modelling and simulation software. The lab activity design was similar in all sections. Teaching materials, resources, and details of the CMB activity are available in CourseSource (Bergan-Roller et al., 2017). Since there were

Table 1. Study population and sample.

Course Instructor	Population of enrolled students, *N*	Consenting students, *n*	Final sample, *n*
A	151	80	38
B	432	382	215
C	243	210	109
Total	826	672	362

no introduced differences in instruction and learning activities, we hypothesised that instructor effects were minimal and negligible to be included in our analyses. Results of a multivariate analysis of variance (MANOVA) indicated an overall non-significant difference ($F(4,718) = 1.25$, $p = 0.29$) for students' pre-test and post-test examined together by course instructor. Table 1 shows that the final sample includes 44% of the entire population of recruited students. The final sample ($n = 362$) was selected from the group of consenting students based on the completeness of collected data. Thus, the data set used in this study had no missing data.

Computational model-based (CMB) activity implementation

The CMB activity was implemented in Week 7 of a regular semester to achieve the objectives of describing the processes of cellular respiration and explaining its dynamic properties (Figure 1). To familiarise students with the modelling functions of Cell Collective, they completed a training activity in Week 6 using direct instruction. Teaching assistants demonstrated how to access and use Cell Collective, including adding or modifying components and relationships to a computational model. Students were also shown examples of how to report and interpret the results of a computational model. After completing the training activity, the cellular respiration activity was introduced to students, and they were asked to complete a pre-test.

In this activity, students first explored a scientific model of cellular respiration and its underlying component processes such as glycolysis, lactic acid fermentation, and the citric acid cycle. Students used their understanding of the model to make predictions about phenomena to be investigated in the activity using computational models. For

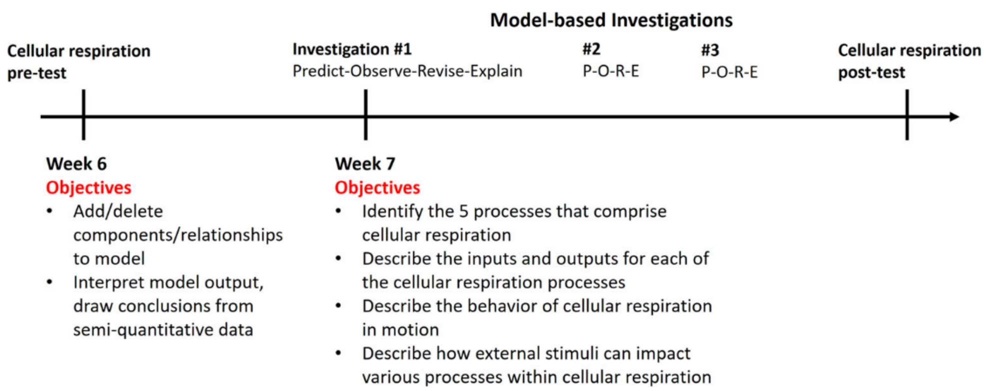


Figure 1. Timeline of lesson activities and objectives. The CMB activity included three model-based investigations providing students with multiple opportunities to iterate the *predict-observe-revise-explain* (PORE) sequence in their modelling practice.

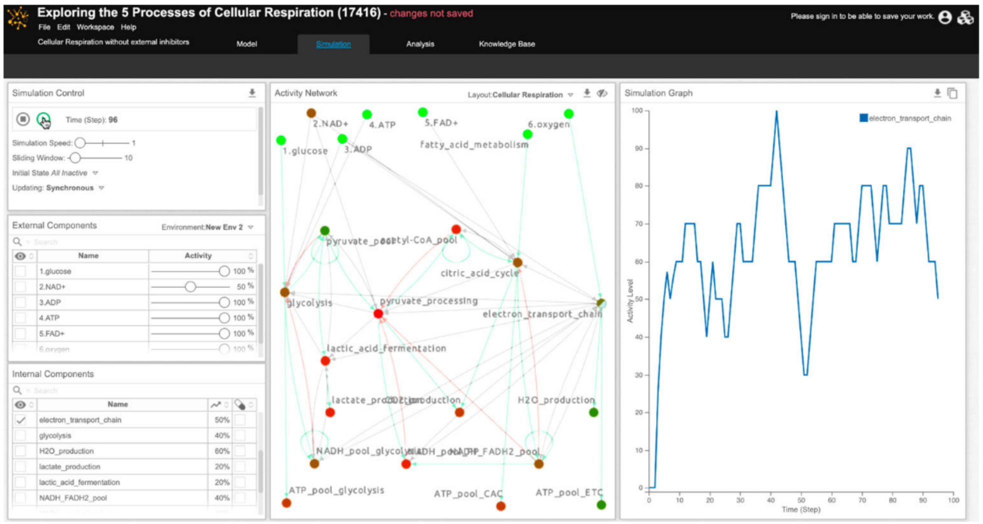


Figure 2. Simulation controls in the CMB activity for cell respiration. A simulation graph for a selected process of cellular respiration is generated after running a computational model.

instance, students were asked to investigate cases such as cancer cells damaging the mitochondria and the use of a drug that inhibits the enzyme that releases electrons into the electron transport chain. Students manipulated the simulation by selecting entities to be included in a cellular respiration model and controlling the activity level of entities involved in the processes of cellular respiration. After setting up a computational model, they ran their simulation and observed a simulation graph of the activity levels of specified cellular respiration processes (Figure 2).

Data

Demographic variables

Demographic data were collected for consenting students from the university registrar's office using unique student identification numbers and signed consent forms. The following data were collected: age, sex, degree-major, ethnicity, first-generation status, class level, and ACT composite score (ACTC). The ethnicity variable had two categories: (a) White, Non-Hispanic, and (b) Non-White. Students who did not identify as White were categorised as Non-White. The variable *degree-major* catalogued students into three groups: (a) declared-science, (b) declared-non-science, (c) undeclared. Undeclared students were those who were still deciding upon a major, transitioning between majors, or working on meeting the requirements to declare a desired major. Students were considered science majors if they had a declared degree in the fields of environmental and life sciences, agriculture, technology, engineering, and physical sciences. Students were categorised as non-science majors if they had a declared degree in the arts and human sciences. Class level referred to students' current standing in the university from first (freshmen) to fourth (senior) year.

Nearly two-thirds (62%) of the participants were female, and ages ranged from 17 to 29 ($M = 18.64$, $SD = 1.23$, Table 2). The majority of the students self-identified as Non-

Table 2. Characteristics of study participants, n = 362.

Variable		*n (%)*
Sex	Male	138 (38.12)
	Female	224 (61.88)
Age, *M* ± *SD*		18.64 ± 1.23
Ethnicity	White, Non-Hispanic	311 (85.91)
	Non-White	51 (14.09)
First-generation status	Yes	108 (29.83)
	No	254 (70.17)
ACTC, *M* ± *SD*		25.30 ± 4.05
Class level	Freshman	255 (70.44)
	Sophomore	66 (18.23)
	Junior/Senior	41 (11.33)
Degree-major	Declared-Science	205 (56.63)
	Declared-Non-Science	42 (11.60)
	Undeclared	115 (31.77)

Note. *M* and *SD* are used to represent mean and standard deviation, respectively.

Hispanic White (86%). About 30% of the students reported that they were the first person in their family to go to college. The ACT composite scores (ACTC) of the whole group ranged from 12 to 35 ($M = 25.30$, $SD = 4.05$). The majority of the students were freshmen (70%) and had declared to pursue a bachelor's degree in a science discipline (57%).

Cellular respiration pre-test and post-test

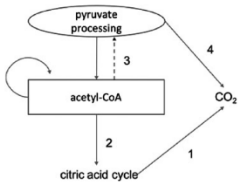
The conceptual test was designed to measure students' knowledge about cellular respiration before and after the computational modelling activity was implemented as a lab activity. Test items included four factual items in a multiple-choice format and an explanation task in the form of a constructed response item. The content of the test was directly tied to concepts addressed in the CMB activity. Items intensified in rigour from easy to difficult questions, that is, from identifying an entity in the cellular respiration system to explaining how the processes in the system are reliant on each other.

The multiple-choice items were given one point each. The constructed-response item was given two points based on a satisfactory mechanistic explanation (Machamer et al., 2000), one point for a partially correct response, and zero for an incorrect response or no response at all (Table 3). Students' responses in *explanation* tasks in the CMB activity are scored in the same way. For the constructed-response item, research assistants were trained to assess entities and activities from mechanisms described by students. About 20% of the 724 student-constructed responses from the pre- and post-test were co-coded by two research assistants and the first author. Coding discrepancies were resolved by consensus. Krippendorff's alpha test (Krippendorff, 2011) was used to estimate the inter-rater reliability. After determining that the inter-rater reliability was acceptable ($\alpha = 0.93$), the remaining 80% of the responses were divided among the three coders for independent scoring.

Cognitive tasks

The PORE tasks were integrated into three model-based investigations comprising the CMB activity. The maximum CMB activity score based on the satisfactory completion of POE tasks in three model-based investigations was 14 points. Raw scores were

Table 3. Scoring example for student explanations in CMB activity and conceptual test.

Why did the carbon dioxide levels decrease to zero when the drug was added?	Incorrect (No point)	Partially correct (1 point)	Correct (2 points)
Model needed to construct a correct explanation: 	No response or incorrect	Correct but incomplete e.g. If the citric acid cycle was inhibited by the drug, (1) the cell would no longer produce CO_2.	Correct entities and activities e.g. If the citric acid cycle was inhibited by the drug, (1) the cell would no longer produce CO_2. Also, (2) acetyl-CoA would accumulate and (3) inhibit pyruvate processing. Therefore, pyruvate processing would also be inactive and (4) not produce CO_2.

converted to percentage scores for ease of comparison. Revision work was used as a grouping variable to investigate variations in learning outcomes.

Prediction scores were derived from student responses to four objective type tasks in the CMB activity prompting them to predict how one variable might change when another variable was manipulated in the interconnected processes of cellular respiration. Students responded to these *prediction* prompts prior to manipulating and running their computational models. Students were provided with a diagram to guide their prediction and subsequent modelling in Cell Collective. The diagram represents a partially worked-out model (Mulder et al., 2016) of unknown entities that interact in an identified cellular respiration phenomenon. After students test their predictions by manipulating a computational model using the cellular respiration simulation, they record the output of their computational models in response to four objective type *observation* tasks.

The CMB activity included prompts for students to evaluate their predictions and observations and write changes based on the results of their simulation of the computational model. Instead of deleting their initial responses to *prediction* and *observation* tasks, changes were recorded as *revisions*. *Revision* work was unique for each student based on the accuracy of their predictions and observations in each of the three investigations. Specifically, students who made correct predictions and observations were not expected to revise their work. Among the group of students who were required to revise, some students did not complete the task. Based on students' *revision* work, three groups of students were identified: (a) not expected to revise (NR), (b) required: revised (RR), and (c) required: did not revise (RDNR) (Figure 3).

The *explanation* tasks at the end of each investigation required students to construct a mechanistic response demonstrating their understanding of entities and activities in the cellular respiration processes involved in the phenomena they investigated through computational modelling. Responses were given a score ranging from 0 to 2 corresponding to incorrect, partially correct, and correct explanations (Table 3).

Analytic approach

To address research question 1, we analysed the impact of student revision on learning gains using binary logistic regression on a smaller sample ($n = 212$) after applying

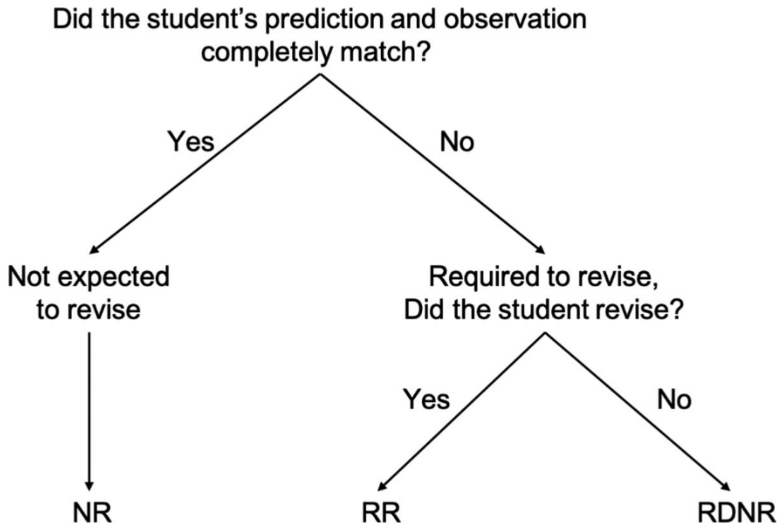


Figure 3. Student groups based on revision work. All students engaged in error detection by comparing predictions and observations.

appropriate data filters. Students were grouped into two categories: positive (gaingroup $= 1$) or negative learning gain (gaingroup $= 0$) from pre- to post-test. Students who had a zero learning gain ($n = 14$) were removed from this analysis to aid in the interpretability of results. While it is intuitive to label a positive gain as an ideal learning outcome and a negative gain as a problematic one, a zero learning gain cannot be interpreted in a straightforward manner. Specifically, a zero learning gain was ideal when a student maintained a relatively high score from pre- to post-test but not for a student with a low score on both pre- and post-test. Students who were not expected to revise (NR) ($n = 109$) were also removed from this analysis to focus on investigating the effect of revision on students who were required to resolve cognitive conflicts between their predictions and computational models. Lastly, we removed students who had a zero or a perfect score ($n = 27$) on the pre-test. These scores create a ceiling that results in measurement inaccuracy for this analysis because the range of possible learning gains is not inclusive of values beyond those points. The probability that a student with a perfect pre-test score would get a positive learning gain is zero. Similarly, the probability that a student with a zero score in the pre-test would get a negative learning gain is also zero. In effect, a smaller sample was used for this analysis, which only included students who could potentially get a negative or positive learning gain upon completion of the cognitive tasks in the CMB activity.

The purpose of this analysis was to understand how student revision in the CMB activity impacted student learning based on the probability of demonstrating a positive learning gain or performing better in the post-test on cellular respiration compared to a baseline score. Since the outcome was a categorical variable, a binomial logistic regression was used to predict the probability of having a positive or negative learning gain in the cellular respiration test given a student's performance in the CMB activity's cognitive tasks. Scores in each of the cognitive tasks were combined to generate a total activity score to represent students' performance in the CMB activity. The odds ratio and predicted probabilities of having a positive versus negative learning gain were calculated for RR and RDNR students.

To address research question 2, we evaluated the impact of the specific cognitive tasks in the CMB activity using the whole sample ($n = 362$) and distinguished them from the impact of student characteristics using hierarchical linear regression. This method allowed the independent variables to be entered sequentially based on theoretical foundations related to research questions. Step 1 included the pre-test score as the baseline of students' background knowledge about cellular respiration. We used students' pre-test performance as a covariate. Step 2 included demographic variables that represented incoming student differences, specifically sex, ethnicity, ACTC, class standing, and declared degree-major. Age was excluded in the set of potential predictors since it was not statistically correlated with the criterion. Step 3 included students' performance in each cognitive task activated by the computational modelling activity. To choose the best subset from the list of potential demographic variables in Step 2, we used stepwise regression with forward and backward elimination. Variables were included in the final set of models if they significantly improved the model fit (probability of F for change in R^2<0.05). Step 3 variables were included successively to examine how each of the activity's cognitive tasks contributed to explaining students' performance in the post-test. Residual analysis was performed to check potential violations of the assumptions of linear regression. Semi-partial correlation coefficients were reported to assess the effect of individual predictors in lieu of standardised beta weights since only the continuous predictor variables were standardised with a mean of zero and standard deviation of one to aid in the interpretability of model results. To compare the final model's predictive power in applying hierarchical linear regression, we also extracted a maximal model using all the potential predictors by Akaike Information Criterion (AIC) in a stepwise algorithm. Statistical analyses were conducted using packages in R (R Core Team, 2019).

Results

Overall, students' performance in the cellular respiration assessment improved from pre- to post-test. Students who engaged in *revision* (RR) were more than twice as likely to demonstrate a positive learning gain compared to students who did not revise despite being guided to do so (RDNR). Post-test scores were associated with students' performance in the cognitive tasks after controlling for student differences.

Student performance in the CMB activity

Students' performance increased from a mean pre-test percent score of 40.5 to a post-test score of 70.0 (Table 4). Their average scores on cognitive tasks were in the middle range: *prediction*, 42.0; *observation*, 66.4; and *explanation*, 49.0 (Table 4).

We hypothesised that students' performance in the cognitive tasks were likely to be correlated with each other and may each have a varying degree of association with students' post-test performance. Table 4 shows that students' *prediction* performance had a stronger positive association with *observation* performance ($r = .28$, $p < .01$) than *explanation* performance ($r = .14$, $p < .01$). Correspondingly, students' performance in the observation tasks was strongly associated with their performance in the *explanation*

Table 4. Means, standard deviations, and correlations with confidence intervals.

Variable	*M*	*SD*	1	2	3	4	5	6
1. Age	18.6	1.2						
2. ACTC	25.3	4.1	-.05 [-.15, .05]					
3. Pre-test	40.5	19.5	.06 [-.05, .16]	.31** [.21, .40]				
4. Prediction	42.0	26.6	.03 [-.08, .13]	.12* [.02, .22]	.11* [.00, .21]			
5. Observation	66.4	29.9	-.01 [-.11, .10]	.21** [.11, .31]	.17** [.06, .26]	.28** [.18, .37]		
6. Explanation	49.0	32.9	.01 [-.09, .12]	.24** [.14, .33]	.14** [.04, .24]	.14** [.04, .24]	.41** [.32, .49]	
7. Post-test	70.0	22.2	-.03 [-.13, .08]	.32** [.23, .41]	.20** [.09, .29]	.19** [.09, .29]	.27** [.17, .36]	.43** [.35, .51]

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

tasks ($r = .41$, $p < .01$). Out of these three cognitive activities, students' performance in the *explanation* tasks had the strongest correlation with post-test performance ($r = .43$, $p < .01$).

Revision and learning gains

To examine the effect of *revision* and other cognitive tasks on learning gains in the CMB activity, a binomial logistic regression was used to predict the probability of a student having a positive learning gain upon completion of the CMB activity. RDNR students had the lowest learning gain ($M = 22.3$, $SD = 25.9$) compared to RR ($M = 29.5$, $SD = 27.5$) and NR ($M = 34.6$, $SD = 24.4$) students (Table 5). RDNR and RR students had a relatively similar mean pre-test performance (RDNR: $M = 39.0$, $SD = 16.2$; RR: $M = 38.5$, $SD = 18.7$) suggesting a similar baseline knowledge in cellular respiration but RR students outperformed RDNR students in the post-test (RDNR: $M = 61.3$, $SD = 22.3$; $RR = 68.0$, $SD = 22.2$).

Although the mean learning gains in Table 5 indicate that students who engaged in *revision* work performed better on the post-test than those who did not revise, at the individual level, students may have a positive or negative learning gain. The binomial logistic regression model showed that students' performance in the POE tasks in the CMB activity ($B = 0.04$, $p < .01$) and their *revision* work ($B = 0.90$, $p < .05$) were both significantly associated with the log odds of a positive learning gain (Figure 4). The predicted probabilities plotted in Figure 4 show that the advantage of RR students compared to RDNR students varied depending on their performance in the POE tasks in the CMB

Table 5. Means and standard deviations for pretest, posttest, and raw learning gains across students' revision work.

		Pre-test		Post-test		Gains	
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*
Not expected (NR)	109	44.8	22.1	79.4	18.8	34.6	24.4
Did not revise (RDNR)	74	39.0	16.2	61.3	22.3	22.3	25.9
Revised (RR)	179	38.5	18.7	68.0	22.2	29.5	27.5

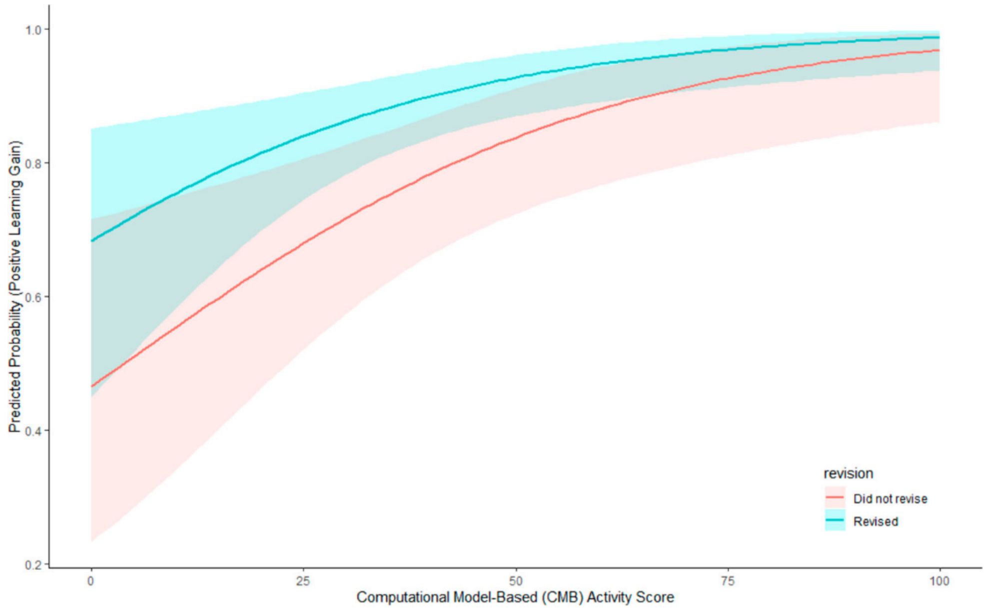


Figure 4. Predicted probabilities of a positive learning gain. The plot was generated from the results of a binomial logistic regression model using the Computational Model-Based (CMB) activity score and *revision* as predictors of demonstrating a positive, raw learning gain. Ribbons indicate 95% confidence intervals.

activity. A likelihood ratio test indicated that this model (Nagelkerke $R^2 = .12$) fits significantly better than an empty model ($\chi^2(2) = 14.58$, $p < .001$).

The exponentiated coefficients (Table 6) show that a one-point increase in score in the CMB activity increases the odds of having a positive learning gain by a factor of 1.04. *Revision* work increases the odds by 2.47; that is, the odds of demonstrating a positive learning gain is more than two times higher for RR students than RDNR students. The result of this analysis does not apply to students with zero learning gain because only the students with positive and negative learning gain were compared. This analysis is limited and does not tell us the impact of revision for all students. Thus, we offer a second analysis in the next section which is a more comprehensive model that includes the whole sample and takes student differences into account.

Predictors of post-test performance

A hierarchical regression analysis was performed to examine the effect of students' performance in cognitive tasks in the CMB activity on their post-test performance after controlling for student differences. The initial model (Model 1) showed pre-test performance to be a significant predictor of post-test performance (adj. $R^2 = .04$; Table 7, Table A1).

Table 6. Logistic regression predicting positive learning gain from CMB activity score and revision.

Predictor	B	Wald χ^2	p	Odds Ratio
CMB activity score	0.04	8.90	.03	1.04
Revision (RR vs. RDNR)	0.90	4.70	<.01	2.47

Table 7. Results of hierarchical linear regression analyses using cellular respiration post-test scores as the criterion, n = 362.

	Model 1			Model 2			Model 6†			Final Model		
	B	SE_B	*r*	*B*	SE_B	*sr*	*B*	SE_B	*sr*	*B*	SE_B	*sr*
Intercept	70.03	1.15		73.22	1.54		75.70	2.68		76.02	2.19	
Step 1: *Baseline Knowledge*												
Pre-test‡	4.34***	1.15	.20	1.85	1.54	.10	1.10	1.07	.00			
Step 2: *Demographic Variables*												
Ethnicity: Non White				−6.02	3.22	.10	−5.60	2.96	.10	−5.43	2.95	.10
ACTC‡				5.92***	1.20	.24	3.81***	1.13	.14	4.11***	1.09	.17
Degree: Non Science				−12.64***	3.53	-.17	−11.61***	3.24	-.17	−12.06***	3.20	-.17
Degree: Undeclared				−2.75	2.48	.00	−2.25	2.28	.00	−2.54	2.26	.00
Step 3: *Cognitive Tasks*												
Prediction‡							1.97	1.05	.10	2.04*	1.03	.10
Observation‡							0.11	1.54	.00			
Revision: Did not revise							−7.67*	3.72	-.10	−7.90*	3.07	-.10
Revision: Revised							−2.54	3.34	.00	−2.84	2.50	.00
Explanation‡							7.03***	1.12	.28	7.10***	1.09	.28
Fit												
R^2	.04				.15			.30			.30	
Adj. R^2	.04				.14			.28			.28	
F for ΔR^2	14.25***				14.55***			39.34***			0.54	

Notes. *B*: regression weight; SE_B: standard error; *sr*: semipartial correlation. †Models 3–5 are included in supplemental materials (Table S1). ‡Standardised with a mean set equal to zero and standard deviation set equal to one. Reference categories: Ethnicity: White, Non-Hispanic, Degree-Major: Science, Revision: Not required.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

When demographic variables were added in Model 2 using forward and backward elimination, ACTC, ethnicity, and degree-major significantly improved model fit (F (4,356) = 14.55, $p < .001$, adj. $R^2 = .14$). In Models 3–6, the cognitive tasks were included in the same order that students engaged in them during the CMB activity (Supplementary Material, Table A1). Specifically, the temporal order of modelling activities in each investigation in the CMB activity followed the PORE sequence and resulted in a statistically significant improvement in model fit with each addition.

Model 6 included all the demographic factors and students' performance in cognitive tasks in the previous models and added *explanation* as a predictor. Model 6 had an adjusted R^2 of .28, which was a significant improvement in model fit compared to Model 5 ($F(1, 351) = 39.34$, $p < .001$). This model indicated that students' performance on the *explanation* tasks was the strongest predictor of their performance on the post-test assessment ($sr = .28$, $p < .001$). *Revision* remained a statistically significant predictor after the performance on the *explanation* tasks was added in the model. Additionally, performance in the *explanation* tasks was lowest among RDNR students ($M = 34.7$, $SD = 29.0$) compared to RR ($M = 44.0$, $SD = 32.1$) and NR ($M = 67.0$, $SD = 28.8$) students (Supplementary Material, Table A2).

Model selection

We compared the full model (Model 6) with a best-subset model generated from stepwise regression. Model 6 and the best-subset model did not differ in terms of model fit (F (2,351) = 0.54, $p = .59$) but Model 6 included two predictors that lacked explanatory power (i.e. pre-test and observation) and therefore added unnecessary complexity. Without these two predictors, the best-subset model indicated that the *prediction* task was statistically significant ($p < .05$). This result was consistent with the hierarchical regression models where the contribution of *prediction* to model fit was stable even with the addition of other cognitive tasks. Model 6 was more complex and failed to improve the fit to the data to justify the increased complexity; therefore, we chose the parsimonious, best-subset model as the final result of our regression analysis.

Final Model. The final model of post-test performance ($R^2 = .30$) included ethnicity ($sr = .10$, $p = .07$), ACTC ($sr = .17$, $p < .001$), degree-major (declared: non-science, $sr = -.17$, $p < .001$; undeclared, $sr = .00$, $p = .26$), *prediction* ($sr = .10$ $p < .05$), *revision* (RDNR, $sr = -.10$, $p < .01$; R, $sr = .00$, $p = .26$) and *explanation* ($sr = .28$, $p < .001$). Among the cognitive tasks in the CMB activity, students' performance on the *explanation* tasks was the strongest predictor of performance on the post-test. RDNR students had the lowest post-test performance. Residual analysis showed that the regression results are valid. The residuals followed a normal distribution when plotted in a histogram and there was an independence of residuals as assessed by the Durbin-Watson statistic of 2.01 ($p = .95$).

Discussion

The implementation of the PORE sequence of cognitive tasks to provide a learning scaffolding for the CMB activity demonstrated that *revision* was strongly associated with learning outcomes. This result provides empirical support for the importance of *revision* in modelling to help students develop increasingly sophisticated explanatory models of complex biological systems (Louca & Zacharia, 2012; Schwarz et al., 2009; Seel, 2017).

Students who revise perform better

In Research Question 1, we asked: '(1) Does revision work influence students' learning gain?' *Revision* was associated with higher odds of demonstrating a positive learning gain. The odds ratio for *revision* indicates that holding POE score constant, RR students are 2.47 times more likely to demonstrate a positive learning gain in the cellular respiration post-test than RDNR students. Examining learning gains is important because students may have different baselines, and learning gains represent students' change in ability and knowledge targeted by the lesson's objectives (Rogaten et al., 2019). RR and RDNR students had similar mean pre-test scores, which suggests that the degree of prior knowledge they possessed about cellular respiration was likely the same prior to the CMB activity. In the pre-test, students were asked to predict how inhibiting the electron transport chain would affect the citric acid cycle and the common responses were typically incorrect descriptions of entities and activities (e.g. 'When you stop the electron transport chain, citric acid cycle would stop because there would be no electrons to transport to the citric acid cycle.') Students also reasoned that because the processes were interconnected, the inhibition of one process halts the other processes. These responses exemplified the difficulties encountered by students when explaining biological phenomena that are intangible and abstract. After completing the CMB activity, the mean post-test score of the RR group was higher than the RDNR group and more students from the RR group (90%) had a positive learning gain compared to the RDNR group (78%). RR students benefited the most from the PORE scaffolding embedded in the CMB activity.

The majority of students in our sample were in the RR group (49%) and they likely represent a typical learner beginning to understand the interactive processes of cellular respiration compared to the NR group (30%) who appeared to have advanced knowledge about cellular respiration. Also, the RR group was the only group that engaged in revision tasks although all groups interacted with error detection prompts after their first observation of the output from their computational model simulations. RR and RDNR students shared the common factor of having incorrect or incomplete models of cellular respiration after they engaged in making predictions and recording observations. However, for the RR group, improvement in learning gains suggests metacognitive affordances of engaging in revision work.

Empirical results in this study support the theoretical body of work emphasising the role of revision in the modelling process; students should be guided to compare observed outcomes of simulated models with expected results as a developmental step toward error detection in prior knowledge and subsequent reconciliation of competing cognitive structures (Fuhrmann et al., 2018; Hewson & Hewson, 1984; Wijnen et al., 2015). Although learning a complex biological system, which involves a vast amount of information, tends to overwhelm biology students (NRC, 2003; Seymour & Hewitt, 2000), iterative revisions during cycles of modelling can be used to manage the complexity to develop an increasingly sophisticated understanding of scientific phenomena (Dauer et al., 2013; Hmelo-Silver et al., 2007; Schwarz et al., 2009).

The impact of *revision* on conceptual change can be attributed to how RR and RDNR students learned differently within the CMB activity. Specifically, the difference in learning gains between RR and RDNR students highlighted the recognisable learning advantage of PORE over POE for students. Explicitly prompting students to engage in error detection did not lead to revision for the RDNR group, which comprised one-third of the sample of students (29%). They chose to short-cut their investigation by directly responding to *explanation* tasks. RDNR students may have tended to make quick interpretations without sufficient evaluation and focused on generating required products, a practice that has been observed in previous studies (Klahr, 2000; Krajcik et al., 1998). They may have also learned to 'game' the system by completing each investigation with minimum effort (Baker et al., 2008). Additionally, *revision* is typically implied in the POE sequence of tasks when used as a learning strategy; that is, students are expected to notice inconsistencies between their initial predictions and observations. However, in our implementation of PORE to provide structure for a CMB activity, results suggest that *revision* may not be a natural strategy for students. Spending time to rework prior learning products might be viewed as extra work (Kallick & Zmuda, 2017). *Revision* should be made an explicit step and a required product in the modelling process because the RR group outperformed the RDNR group in terms of learning gains.

Students' performance in prediction, revision, and explanation tasks predict learning

In Research Question 2, we asked: 'To what extent do cognitive tasks in a CMB activity applied in a *predict-observe-revise-explain* (PORE) sequence predict student post-test performance in cellular respiration after controlling for student differences?' Students' performance in *prediction*, *revision*, and *explanation* tasks were significantly associated with post-test performance after accounting for student differences. The final regression model accounted for ethnicity, ACTC, and degree-major due to their unique contribution to the amount of variance explained by the model; ACTC and degree-major were significantly associated with student learning. Because ACTC reflects general high school knowledge level, the model may be improved with the inclusion of the type and extent of cellular respiration instruction that students experienced in high school to account for student differences. Prior studies in K-12 biology instruction reveal that students hold many misconceptions about the processes of cellular respiration (Driver et al., 2014; Flores et al., 2003; Songer & Mintzes, 1994) but exposure to reform-based instruction has shown success in improving high school students' conceptual knowledge of cellular respiration (Brown, 2003; Çakir et al., 2002; Dam et al., 2019).

Prediction

Students' performance in *prediction* tasks was significantly associated with performance in the post-test, although the effect was small after including *revision* and *explanation* as predictors of learning. This result indicates that incorrect predictions could be productive for students because they lay the groundwork for cognitive conflict and subsequent revision needed to improve their understanding of a complex system such as cellular respiration. Making predictions and creating a hypothesis about a model were previously shown to help students evaluate models effectively and perform better in a subsequent test of conceptual knowledge (Löhner et al., 2005). As students became cognitively aware of what they initially knew about cellular respiration, they were primed to compare their predictions with the results of a computational investigation. *Prediction* was weakly associated with students' performance in the *explanation* tasks which is likely due to change in students' understanding shaped by model observations and revisions in the CMB activity (Louca et al., 2011; Pennington et al., 2016). Students who make incorrect predictions, and they are likely to do so due to the complex nature of cellular respiration, may still perform well in constructing scientific explanations depending on their performance in other cognitive tasks in the CMB activity.

Observation

Because students' conception of cellular respiration processes was developing as they work through the CMB activity, it was expected that their predictions and observations would present cognitive conflicts (Posner et al., 1982). In the *observation* phase, the CMB activity prescribed the variables to observe and report and therefore left little room to consider alternatives, deliberate why they were reporting certain values, or actively interpret the results. Thus, students' responses to the observation prompts did not reflect what they learned regarding the relationships between processes in the system. This could also explain why students' performance in the *observation* tasks was not significantly associated with post-test results. Compared to other tasks in PORE, the *observation* tasks in the CMB activity were easier to accomplish because students simply reported the output of their computational models. Nonetheless, students' performance during the *observation* tasks remained an integral part of the learning process to develop students' understanding of the interactive processes of cellular respiration. The strong association between students' performance in the *observation* and *explanation* tasks indicates that what students observe in their investigations with computational models is associated with the quality of their scientific explanations.

Revision

An important finding of the study is that RR students performed just as well as NR students in the post-test. Students who encountered cognitive conflicts in the modelling process likely improved their conceptual understanding of cellular respiration system dynamics through engagement in *revision* tasks. Although each cognitive task in the PORE sequence could have an indirect effect on the post-test through their association with the next task, adding *revision* was particularly important to facilitate student learning.

When students engage in *revision*, they demonstrate retrospective action (Xenofontos et al., 2019). Returning to previous stages of a learning activity sequence enhances student learning especially among students who did not devote adequate time to working on certain tasks on their first pass (Xenofontos et al., 2019). The *revision* phase was the students' second attempt to make predictions and record observations to resolve their cognitive conflicts. Model-based activities can be improved by providing opportunities for students to return to their initial models and modify them instead of focusing only on model generation which remains pervasive in modelling instruction (Abrahamson et al., 2007; Soderberg & Price, 2003). Additionally, when RR students reasoned about cellular respiration after engaging in a *predict-observe-revise* sequence of cognitive tasks, they performed better in *explanation* tasks and in the post-test than

RDNR students. RDNR students' poorer performance in explaining cellular respiration system dynamics and retrieving key concepts needed to demonstrate knowledge in the post-test reflects a lack of development of deeper understanding.

Explanation

Students' performance in the *explanation* tasks was the strongest predictor of their performance on the post-test because it required processing of knowledge gained from collective learning experiences following the PORE sequence. Students' performance in each of the cognitive tasks completed prior to constructing a scientific explanation was significantly correlated with their performance in the *explanation* tasks, respectively. These correlations suggest that explanations provided by students were likely model-based, drawing from their predictions and observations of their computational models. The PORE tasks in the CMB activity guided students to construct a model-based explanation of cellular respiration system dynamics as opposed to using an experiment to confirm an instructor-provided scientific explanation. This is important because guiding students through cycles of modelling to build their own explanations for scientific phenomena instead of learning explanations brings authenticity to science learning (Hester et al., 2018). When students explained phenomena in cellular respiration in the CMB activity, responses to the *explanation* tasks showed that they drew from knowledge gained from a complex interaction between their simulation data and computational model characterising modelling-based learning (Hester et al., 2018; Passmore et al., 2009; Wilensky & Rand, 2015).

We also found that students' performance in *explanation* tasks was strongly associated with students' performance on *observation* and *revision* tasks. This means that students can be supported to successfully construct scientific explanations through investigations and opportunities to reflect upon initial predictions and to revise models. Because students are likely to make incorrect predictions about complex biological systems such as cellular respiration, the CMB activity had the potential to support students to transition toward a more sophisticated understanding of the system through engagement in cognitive tasks that characterised practices scientists use in explaining the behaviours of complex phenomena (de Jong, 2006; de Jong & van Joolingen, 1998; Wilensky & Rand, 2015). Scientists constantly modify their conceptions that are found to be inconsistent with empirical evidence and use revised models for communicating and explaining their findings so that they may be further scrutinised and updated by the scientific community (Nersessian, 2009). When students are guided through a similar process and learn to use models to explain phenomena in an observed biological system, their conceptual understanding becomes more nuanced and the modelling experience refines their science literacy (Ke et al., 2021).

Overall, the use of the PORE sequence of cognitive tasks in the CMB activity appeared to support conceptual change in cellular respiration. The process leading to students' construction of scientific explanations in the CMB activity reflects how students learn in science through modelling (Gilbert, 1991; Lehrer & Schauble, 2000; Windschitl et al., 2008). Providing a cognitive scaffolding for the CMB activity also emphasised the learning of science concepts through the process of doing science as opposed to traditional curricula designed to teach scientific skills and concepts separately. Previous studies have shown that the lack of integration of scientific practices in learning science has resulted in inaccurate views about how scientific knowledge is generated and overemphasis on a vast amount of fragmented science facts that students report as difficult to learn (Hester et al., 2018; McComas & Kampourakis, 2015; Southard et al., 2016). The PORE sequence emphasised an explanation-oriented approach which cognitively activated students to formulate predictions, analyse and interpret computational data patterns, connect observational data with theoretical concepts, revise their model, and construct an evidence-based explanation of a scientific phenomenon.

Implications for teaching and future research

Results of this study can be used to guide the development of CMB activities for the learning of complex biological systems. Cognitive scaffolding for CMB activities that can increase engagement in *revision* is necessary to be developed. The present study showed that even with the inclusion of explicit revision prompts such as the use of the PORE structure for the CMB activity, students may still resort to quick interpretations and favour minimum effort. This implies the need to identify which students choose to engage in the revision process and why students may ignore inconsistencies between predictions and observations.

Research in cognitive science about how people learn emphasises the necessity of scaffolding learning situated in one context to facilitate transferability in solving problems in another context (NRC, 2000a; Quellmalz et al., 2016). Using the PORE sequence in CMB activities can provide students with a cognitive scaffolding that may be transferable to other scientific investigations that integrate practices and content.

It is also important to note that in the CMB activity used in this study, the iterative process of engaging in PORE tasks was conducted using three successive investigations about processes in cellular respiration. The repetition of the use of PORE tasks in these investigations may have contributed to learning outcomes because it resembled the frequent use of cognitive tools displayed by scientists when reasoning with models. Examining the impact of deliberate practice would be worthwhile because students can improve in making predictions, observing, revising, and explaining, due to the repetition of cognitive tasks in each model-based investigation. Additionally, the PORE tasks can be used by educators to track student thinking through their written responses to guide questions within each task. These tasks result in learning products that can provide valuable insight for monitoring student learning.

Acknowledgements

We extend our gratitude to the undergraduate students, graduate student teaching assistants, and professors who supported this work or who took the time to take part in this study. We thank our undergraduate research assistants, Autumn Fluent and Lauren Brickett, for their participation in data preparation, cleaning, and coding. We are also thankful to Dr. Gretchen King and Marius Dongmo for their preliminary analyses on student learning gains.

Disclosure statement

TH is the majority stakeholder in a company, Discovery Collective, Inc., that has proprietary rights to the software used in this project. No promotion of Discovery Collective products to the exclusion of other similar products should be construed.

Funding

This work was supported by the National Science Foundation (NSF) Division of Undergraduate Education: Grant Number DUE 1432001. Ideas presented in this manuscript are those of the authors and do not necessarily reflect the views of personnel affiliated with the NSF.

Ethics statement

This study was approved under the exempt research category by the Institutional Review Board at the University of Nebraska-Lincoln and was conducted in established or commonly accepted educational settings involving normal educational practices.

Notes on contributors

Lyrica Lucas is a Postdoctoral Research Associate in the School of Natural Resources at the University of Nebraska-Lincoln. Dr. Lucas obtained her bachelor's degree in Physics and Technology Education from the Philippine Normal University, and her M.A. in Physics from National Institute of Physics at the University of the Philippines-Diliman. She received her M.A. in Teaching and Learning as a Fulbright Scholar and Ph.D. in Educational Studies with a focus on science education at the University of Nebraska-Lincoln. Her current research investigates the use of quantitative modelling activities in science learning and aspects of science teacher education that facilitate the use of inquiry-based practices in classroom instruction, discourse, and assessment.

Tomáš Helikar is an Associate Professor in the Department of Biochemistry at the University of Nebraska-Lincoln. Dr. Helikar holds courtesy appointments as an Associate Professor in the Departments of Computer Science and Engineering, and Pharmacology and Experimental Neuroscience. Dr. Helikar obtained his B.Sc. in Bioinformatics from the Computer Science Department at the University of Nebraska at Omaha, and his Ph.D. from the Department of Pathology and Microbiology at the University of Nebraska Medical Centre. Dr. Helikar joined the University of Nebraska-Lincoln in 2013 as an Assistant Professor after completing a 3-year postdoctoral work in the Department of Mathematics at the University of Nebraska at Omaha. Dr. Helikar's research program centres around multi-scale modelling of the immune system, and technology development in an effort to make computational modelling accessible to anyone, regardless of their computational background. His group develops, Cell Collective, a modelling and simulation software that is now used by many universities and high schools to teach biology through hands-on modelling.

Joseph Dauer is an Associate Professor of Life Science Education at the University of Nebraska-Lincoln. He has a B.S. in Biology/Mathematics from Western Washington University and M.S. and Ph.D. in Ecology from Penn State University. He conducted post-doctoral research on plant population dynamics and student learning in biology through modelling at Michigan State University. Currently his research focus includes teaching and learning of biology through quantitative modelling and an exploration of neurobiology during modelling activities. He teaches Introductory Biology, Ecology, and College Science Teaching at the University of Nebraska-Lincoln.

ORCID

Lyrica Lucas http://orcid.org/0000-0002-6833-8353
Tomáš Helikar http://orcid.org/0000-0003-3653-1906
Joseph Dauer http://orcid.org/0000-0002-8971-0441

References

Abrahamson, D., Blikstein, P., & Wilensky, U. (2007). Classroom model, model classroom: Computer-supported methodology for investigating collaborative-learning pedagogy. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the computer supported collaborative learning (CSCL) conference* (Vol. 8, Part 1, pp. 46–55). Rutgers University.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

American Association for the Advancement of Science (AAAS). (2011). *Vision and change in undergraduate biology education*. http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research*, *19*(2), 185–224. https://www.learntechlib.org/primary/p/24328/

Bartley, J., Riedel, M., Salo, T., Boeving, E., Bottenhorn, K., Bravo, E., Odean, R., Nazareth, A., Laird, R., Sutherland, M., Pruden, S., Brewe, E., & Laird, A. (2019). Brain activity links performance in science reasoning with conceptual approach. *npj Science of Learning*, *4*(1), 20. https://doi.org/10.1038/s41539-019-0059-8

Bartocci, E., & Lió, P. (2016). Computational modeling, formal analysis, and tools for systems biology. *PLoS Computational Biology*, *12*(1), e1004591. https://doi.org/10.1371/journal.pcbi.1004591

Bergan-Roller, H. E., Galt, N. J., Chizinski, C. J., Helikar, T., & Dauer, J. T. (2018). Simulated computational model lesson improves foundational systems thinking skills and conceptual knowledge in biology students. *BioScience*, *68*(8), 612–621. https://doi.org/10.1093/biosci/biy054

Bergan-Roller, H. E., Galt, N. J., Dauer, J. T., & Helikar, T. (2017). Discovering cellular respiration with computational modeling and simulations. *CourseSource*. https://doi.org/10.24918/cs.2017.10

Bergan-Roller, H. E., Galt, N. J., Helikar, T., & Dauer, J. T. (2020). Using concept maps to characterise cellular respiration knowledge in undergraduate students. *Journal of Biological Education*, *54*(1), 33–46. https://doi.org/10.1080/00219266.2018.1541001

Booth, C. S., Song, C., Howell, M. E., Rasquinha, A., Saska, A., Helikar, R., Sikich, S. M., Couch, B. A., van Dijk, K., Roston, R. L., & Helikar, T. (2021). Teaching metabolism in upper-division undergraduate biochemistry courses using online computational systems and dynamical models improves student performance. *CBE-Life Sciences Education*, *20*(1), ar13 1–16. https://doi.org/10.1187/cbe.20-05-0105

Brodland, G. W. (2015). How computational models can help unlock biological systems. *Seminars in Cell & Developmental Biology*, *47-48*, 62–73. https://doi.org/10.1016/j.semcdb.2015.07.001

Brookman-Byrne, A., Mareschal, D., Tolmie, A. K., & Dumontheil, I. (2018). Inhibitory control and counterintuitive science and maths reasoning in adolescence. *PLoS ONE*, *13*(6), e0198973. https://doi.org/10.1371/journal.pone.0198973

Brown, D. (2003). High school biology: A group approach to concept mapping. *The American Biology Teacher*, *65*(3), 192–197. https://doi.org/10.2307/4451473

Buckley, B. C., Gobert, J. D., Kindfield, A. C. H., Horwitz, P., Tinker, R. F., Gerlits, B., Wilensky, U., Dede, C., & Willett, J. (2004). Model-based teaching and learning with BioLogica™: what do they learn? How do they learn? How do we know? *Journal of Science Education and Technology*, *13*(1), 23–41. https://doi.org/10.1023/B:JOST.0000019636.06814.e3

Çakir, Ö, Geban, Ö, & Yürük, N. (2002). Effectiveness of conceptual change text-oriented instruction on students' understanding of cellular respiration concepts. *Biochemistry and Molecular Biology Education*, *30*(4), 239–243. https://doi.org/10.1002/bmb.2002.494030040095

Cinici, A., & Demir, Y. (2013). Teaching through cooperative POE tasks: A path to conceptual change. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *86*(1), 1–10. https://doi.org/10.1080/00098655.2012.712557

Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, *22*(9), 1041–1053. https://doi.org/10.1080/095006900416901
Coştu, B., Ayas, A., & Niaz, M. (2012). Investigating the effectiveness of a POE-based teaching activity on students' understanding of condensation. *Instructional Science*, *40*(1), 47–67. https://doi.org/10.1007/s11251-011-9169-2
Dam, M., Ottenhof, K., Van Boxtel, C., & Janssen, F. (2019). Understanding cellular respiration through simulation using lego® as a concrete dynamic model. *Education Sciences*, *9*(2), 72. https://doi.org/10.3390/educsci9020072
Dauer, J. T., & Long, T. M. (2015). Long-term conceptual retrieval by college biology majors following model-based instruction. *Journal of Research in Science Teaching*, *52*(8), 1188–1206. https://doi.org/10.1002/tea.21258
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. https://doi.org/10.1002/tea.21094
de Jong, T. (2006). Technological advances in inquiry learning. *Science*, *312*(5773), 532–533. https://doi.org/10.1126/science.1127750
de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179–201. https://doi.org/10.3102/00346543068002179
Doerr, H. M. (1997). Experiment, simulation and analysis: An integrated instructional approach to the concept of force. *International Journal of Science Education*, *19*(3), 265–282. https://doi.org/10.1080/0950069970190302
Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (2014). *Making sense of secondary science: Research into children's ideas 2nd edition*. Routledge.
Flores, F., Tovar, M. E., & Gallegos, L. (2003). Representation of the cell and its processes in high school students: An integrated view. *International Journal of Science Education*, *25*(2), 269–286. https://doi.org/10.1080/09500690210126793
Fugelsang, J., & Dunbar, K. (2005). Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, *43*(8), 1204–1213. https://doi.org/10.1016/j.neuropsychologia.2004.10.012
Fuhrmann, T., Schneider, B., & Blikstein, P. (2018). Should students design or interact with models? Using the bifocal modelling framework to investigate model construction in high school science. *International Journal of Science Education*, *40*(8), 867–893. https://doi.org/10.1080/09500693.2018.1453175
Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching*, *28*(1), 73–79. https://doi.org/10.1002/tea.3660280107
Haysom, J., & Bowen, M. (2010). *Predict, observe, explain: Activities enhancing scientific understanding*. NSTA Press.
Helikar, T., Cutucache, C. E., Dahlquist, L. M., Herek, T., Larson, J., & Rogers, J. A. (2015). Integrating interactive computational modeling in biology curricula. *PLoS Computational Biology*, *11*(3), e1004131. https://doi.org/10.1371/journal.pcbi.1004131
Helikar, T., Kowal, B., McClenathan, S., Bruckner, M., Rowley, T., Madrahimov, A., Wicks, B., Shrestha, M., Limbu, K., & Rogers, J. A. (2012). The cell collective: Toward an open and collaborative approach to systems biology. *BMC Systems Biology*, *6*(1), 96. https://doi.org/10.1186/1752-0509-6-96
Hester, S., Nadler, M., Katcher, J., Elfring, L., Dykstra, E., Rezende, L., & Bolger, M. (2018). Authentic inquiry through modeling in biology (AIM-Bio): An introductory laboratory curriculum that increases undergraduates; scientific agency and skills. *CBE-Life Sciences Education*, *17*(4), ar63 1–23. https://doi.org/10.1187/cbe.18-06-0090
Hewson, P. W., & Hewson, M. G. A. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science*, *13*(1), 1–13. https://doi.org/10.1007/BF00051837
Hmelo-Silver, C. E., 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. https://doi.org/10.1080/10508400701413401

Kallick, B., & Zmuda, A. (2017). *Students at the center: Personalized learning with habits of the mind*. ASCD.
Ke, L., Sadler, T. D., Zangori, L., & Friedrichsen, P. J. (2021). Developing and using multiple models to promote scientific literacy in the context of socio-scientific issues. *Science & Education*, *30*, 589–607. https://doi.org/10.1007/s11191-021-00206-1
King, G. P., Bergan-Roller, H., Galt, N., Helikar, T., & Dauer, J. T. (2019). Modelling activities integrating construction and simulation supported explanatory and evaluative reasoning. *International Journal of Science Education*, *41*(13), 1764–1786. https://doi.org/10.1080/09500693.2019.1640914
Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. The MIT Press.
Krajcik, J., Berger, C. F., & Czerniak, C. M. (2002). *Teaching science in elementary and middle school classrooms: A project-based approach (2nd ed.)*. McGraw Hill.
Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., & Fredricks, J. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, *7*(3&4), 313–350. https://doi.org/10.1080/10508406.1998.9672057
Krajcik, J., & Merritt, J. (2012). Engaging students in scientific practices: What does constructing and revising models look like in the science classroom? *The Science Teacher*, *79*(3), 38–41. https://www.jstor.org/stable/43557386
Krajcik, J., & Mun, K. (2014). Promises and challenges of using learning technologies to promote student learning of science. In N. G. Lederman, & S. K. Abell (Eds.), *Handbook of research on science education: Volume II* (pp. 337–360). Routledge.
Krippendorff, K. (2011). *Computing krippendorff's alpha-reliability*. Annenberg School for Communication Departmental Papers.
Lazonder, A. W., & Harmsen, R. (2016). Meta-Analysis of inquiry-based learning. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366
Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of Applied Developmental Psychology*, *21*(1), 39–48. https://doi.org/10.1016/S0193-3973(99)00049-0
Lesh, R., & Doerr, H. M. (2000). Symbolizing, communicating, and mathematizing: Key components of models and modeling. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms. Perspectives on discourse, tools, and instructional design* (pp. 361–383). Erlbaum.
Liew, C., & Treagust, D. (1998, April 13-17). The effectiveness of predict-observe-explain tasks in diagnosing students' understanding of science and in identifying their levels of achievement. Annual Meeting of the American Educational Research Association, San Diego, CA. https://files.eric.ed.gov/fulltext/ED420715.pdf.
Löhner, S., van Joolingen, W. R., Savelsbergh, E. R., & van Hout-Wolters, B. (2005). Students' reasoning during modeling in an inquiry learning environment. *Computers in Human Behavior*, *21*(3), 441–461. https://doi.org/10.1016/j.chb.2004.10.037
Louca, L., & Zacharia, Z. (2012). Modeling-based learning in science education: Cognitive, metacognitive, social, material and epistemological contributions. *Educational Review*, *64*(4), 471–492. https://doi.org/10.1080/00131911.2011.628748
Louca, L., Zacharia, Z., Michael, M., & Constantinou, C. P. (2011). Objects, entities, behaviors and interactions: A typology of student-constructed computer-based models of physical phenomena. *Journal of Educational Computing Research*, *44*(2), 173–201. https://doi.org/10.2190/EC.44.2.c
Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25. https://doi.org/10.1086/392759

MacLeod, M., & Nersessian, N. J. (2018). Modeling complexity: Cognitive constraints and computational model-building in integrative systems biology. *History and Philosophy of the Life Sciences*, *40*(1), 17. https://doi.org/10.1007/s40656-017-0183-9

Mayes, R., Forrester, J., Christus, J. S., Peterson, F., & Walker, R. (2014). Quantitative reasoning learning progression: The matrix. *Numeracy*, *7*(2), 5. https://doi.org/10.5038/1936-4660.7.2.5

McComas, W. F., & Kampourakis, K. (2015). Using the history of biology, chemistry, geology and physics to teach aspects of the nature of science. *Review of Science, Mathematics and ICT Education*, *9*(1), 47–76. https://doi.org/10.26220/REV.2240

Mulder, Y. G., Bollen, L., de Jong, T., & Lazonder, A. W. (2016). Scaffolding learning by modelling: The effects of partially worked-out models. *Journal of Research in Science Teaching*, *53*(3), 502–523. https://doi.org/10.1002/tea.21260

National Institute of Biomedical Imaging and Bioengineering (NIBIB). (2018). Computational modeling. https://www.nibib.nih.gov/science-education/science-topics/computational-modeling.

National Research Council (NRC). (2000a). *How people learn: Brain, mind, experience, and school*. The National Academies Press.

National Research Council (NRC). (2000b). *Inquiry and the national science education standards: A guide for teaching and learning*. The National Academies Press.

National Research Council (NRC). (2003). *BIO 2010: Transforming undergraduate education for future research biologists*. National Academies Press.

Nenciovici, L., Allaire-Duquette, G., & Masson, S. (2019). Brain activations associated with scientific reasoning: A literature review. *Cognitive Processing*, *20*(2), 139–161. https://doi.org/10.1007/s10339-018-0896-z

Nersessian, N. J. (2009). How Do engineering scientists think? Model-based simulation in biomedical engineering research laboratories. *Topics in Cognitive Science*, *1*(4), 730–757. https://doi.org/10.1111/j.1756-8765.2009.01032.x

Papaevripidou, M., & Zacharia, Z. C. (2015). Examining how students' knowledge of the subject domain affects their process of modeling in a computer programming environment. *Journal of Computers in Education*, *2*(3), 251–282. https://doi.org/10.1007/s40692-015-0034-1

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics*, *109*(7), 394–402. https://doi.org/10.1111/j.1949-8594.2009.tb17870.x

Penner, D. E. (2000). Chapter 1: Cognition, computers, and synthetic science: Building knowledge and meaning through modeling. *Review of Research in Education*, *25*(1), 1–35. https://doi.org/10.3102/0091732X025001001

Pennington, D., Bammer, G., Danielson, A., Gosselin, D., Gouvea, J., Habron, G., Hawthorne, D., Parnell, R., Thompson, K., Vincent, S., & Wei, C. (2016). The EMBeRS project: Employing model-based reasoning in socio-environmental synthesis. *Journal of Environmental Studies and Sciences*, *6*(2), 278–286. https://doi.org/10.1007/s13412-015-0335-8

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes US think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001

Perkins, K., Adams, W., Dubsson, M., Finkelstein, N., Reid, S., & Wieman, C. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher*, *44*(1), 18–23. https://doi.org/10.1119/1.2150754

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chicago Press.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227. https://doi.org/10.1002/sce.3730660207

Qudrat-Ullah, H. (2010). Perceptions of the effectiveness of system dynamics-based interactive learning environments: An empirical study. *Computers & Education*, *55*(3), 1277–1286. https://doi.org/10.1016/j.compedu.2010.05.025

Quellmalz, E., Silberglitt, M., Buckley, B., Loveland, M., & Brenner, D. (2016). Simulations for supporting and assessing science literacy. In Y. Rosen, S. Ferrera, & M. Mosharraf (Eds.), *Handbook of research on technology tools for real-world skill development* (pp. 191–229). IGI Global.

Quintana, C., Reiser, B., Davis, E., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386. https://doi.org/10.1207/s15327809jls1303_4

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rogaten, J., Rientes, B., Sharpe, R., Cross, S., Whitelock, D., Lygo-Baker, S., & Littlejohn, A. (2019). Reviewing affective, behavioural and cognitive learning gains in higher education. *Assessment & Evaluation in Higher Education*, *44*(3), 321–337. https://doi.org/10.1080/02602938.2018.1504277

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education*, *58*(1), 136–153. https://doi.org/10.1016/j.compedu.2011.07.017

Sarma, G., & Faundez, V. (2017). Integrative biological simulation praxis: Considerations from physics, philosophy, and data/model curation practices. *Cellular Logistics*, *7*(4), e1392400. https://doi.org/10.1080/21592799.2017.1392400

Savander-Ranne, C., & Kolari, S. (2003). Promoting the conceptual understanding of engineering students through visualization. *Global Journal of Engineering Education*, *7*(2), 189–199. http://www.wiete.com.au/journals/GJEE/Publish/vol7no2/SavRanneKolari.pdf

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632–654. https://doi.org/10.1002/tea.20311

Seel, N. M. (2017). Model-based learning: A synthesis of theory and research. *Educational Technology Research and Development*, *65*(4), 931–966. https://doi.org/10.1007/s11423-016-9507-9

Seymour, E., & Hewitt, N. (2000). *Talking about leaving: Why undergraduates leave the sciences*. Westview Press.

Sins, P. H. M., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The difficult process of scientific modelling: An analysis of novices' reasoning during computer-based modelling. *International Journal of Science Education*, *27*(14), 1695–1721. https://doi.org/10.1080/09500690500206408

Soderberg, P., & Price, F. (2003). An examination of problem-based teaching and learning in population genetics and evolution using EVOLVE, a computer simulation. *International Journal of Science Education*, *25*(1), 35–55. https://doi.org/10.1080/09500690110095285

Songer, C., & Mintzes, J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, *31*(6), 621–637. https://doi.org/10.1002/tea.3660310605

Southard, K., Wince, T., Meddleton, S., & Bolger, M. S. (2016). Features of knowledge building in biology: Understanding undergraduate students' ideas about molecular mechanisms. *CBE—Life Sciences Education*, *15*(1), ar7 1–16. https://doi.org/10.1187/cbe.15-05-0114.

Svoboda, J., & Passmore, C. (2013). The strategies of modeling in biology education. *Science & Education*, *22*(1), 119–142. https://doi.org/10.1007/s11191-011-9425-5

Sweeney, L. B., & Sterman, J. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *System Dynamics Review*, *23*(2/3), 285–311. https://doi.org/10.1002/sdr.366

Usher, D. C., Driscoll, T. A., Dhurjati, P., Pelesko, J. A., Rossi, L. F., Schleiniger, G., Pusecker, K., & White, H. B. (2010). A transformative model for undergraduate quantitative biology education. *CBE—Life Sciences Education*, *9*(3), 181–188. https://doi.org/10.1187/cbe.10-03-0029

van Joolingen, W., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, *21*(4), 671–688. https://doi.org/10.1016/j.chb.2004.10.039

van Mil, M., Boerwinkel, D. J., & Waarlo, A. J. (2013). Modelling molecular mechanisms: A framework of scientific reasoning to construct molecular-level explanations for cellular behaviour. *Science & Education*, *22*(1), 93–118. https://doi.org/10.1007/s11191-011-9379-7

White, R. T., & Gunstone, R. F. (1992). *Probing understanding*. The Falmer Press.

Wijnen, F. M., Mulder, Y. G., Alessi, S. M., & Bollen, L. (2015). The potential of learning from erroneous models: Comparing three types of model instruction. *System Dynamics Review*, *31*(4), 250–270. https://doi.org/10.1002/sdr.1546

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with NetLogo. Cambridge, massachusetts*. The MIT Press.

Wilkerson, M., Shareff, B., Gravel, B., Shaban, Y., & Laina, V. (2017). Exploring computational modeling environments as tools to structure classroom-level knowledge building. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a difference: Prioritizing equity and access in CSCL, 12th international conference on computer supported collaborative learning (CSCL), volume 1* (pp. 447–454). International Society of the Learning Sciences.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967. https://doi.org/10.1002/sce.20259

Xenofontos, N. A., Hovardas, T., Zacharia, Z. C., & de Jong, T. (2019). Inquiry-based learning and retrospective action: Problematizing student work in a computer-supported learning environment. *Journal of Computer Assisted Learning*, *36*(1), 12–28. https://doi.org/10.1111/jcal.12384

Zacharia, Z. C. (2005). The impact of interactive computer simulations on the nature and quality of postgraduate science teachers' explanations in physics. *International Journal of Science Education*, *27*(14), 1741–1767. https://doi.org/10.1080/09500690500239664

Table A1. Results of multivariate hierarchical regression analyses using cellular respiration post-test scores as the criterion, n = 362.

	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	B	SE_B	*r*	*B*	SE_B	*sr*	*B*	SE_B	*sr*	*B*	SE_B	*sr*	*B*	SE_B	*sr*	*B*	SE_B	*sr*
Intercept	70.03	1.15		73.22	1.54		73.20	1.52		73.14	1.50		77.00	2.82		75.70	2.68	
Step 1: *Baseline Knowledge*																		
Pre-test+	4.34***	1.15	.20	1.85	1.54	.10	1.62	1.15	.00	1.25	1.14	.00	1.30	1.12	.00	1.10	1.07	.00
Step 2: *Demographic Variables*																		
Ethnicity: Non White				−6.02	3.22	-.10	−5.76	3.19	-.10	−5.95	3.14	-.10	−6.58*	3.11	-.10	−5.60	2.96	-.10
ACTC†				5.92***	1.20	.24	5.62***	1.19	.22	5.04***	1.18	.20	4.69***	1.18	.17	3.81***	1.13	.14
Degree: Non Science				−12.64***	3.53	-.17	−12.27***	3.50	-.17	−12.93***	3.44	-.17	−12.59***	3.41	-.17	−11.61***	3.24	-.17
Degree: Undeclared				−2.75	2.48	.00	−2.95	2.46	.00	−2.44	2.42	.00	−2.49	2.40	.00	−2.25	2.28	.00
Step 3: *Cognitive Tasks*																		
Prediction†							3.01**	1.09	.14	1.96	1.11	.10	2.08	1.10	.10	1.97	1.05	.10
Observation†										4.06***	1.13	.17	2.30	1.58	.00	0.11	1.54	.00
Revision: Did not revise													−10.56**	3.89	-.14	−7.67*	3.72	-.10
Revision: Revised													−3.30	3.52	.00	−2.54	3.34	.00
Explanation†																7.03***	1.12	.28
Fit																		
R^2	.04			.15			.17			.20			.22			.30		
Adj. R^2	.04			.14			.16			.19			.20			.28		
F for ΔR^2	14.25***			14.55***			8.97**			14.65***			5.35**			39.34***		

Note. *B*: regression weight; SE_B: standard error; *sr*: semipartial correlation. †Standardised with a mean set equal to zero and standard deviation set equal to one.
Reference categories: Ethnicity: White, Non-Hispanic, Degree-Major: Science, Revision: Not required.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table A2. Means and standard deviations for explanation tasks across students' revision work.

		Explanation	
	n	*M*	*SD*
Not expected (NR)	109	67.0	28.8
Did not revise (RDNR)	74	34.7	29.0
Revised (RR)	179	44.0	32.1