

Molecular Concepts Adaptive Assessment (MCAA) Characterizes Undergraduate Misconceptions about Molecular Emergence

Andrea Gauthier,[†] Stuart Jantzen,[‡] Gaël McGill,[§] and Jodie Jenkinson^{†*}

[†]Institute of Medical Science, University of Toronto, Toronto, ON M5S 1A8, Canada; [‡]Biomedical Communications, Department of Biology, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada; [§]Center for Molecular and Cellular Dynamics, Harvard Medical School, Boston, MA 02115

ABSTRACT

This paper discusses the results of two experiments assessing undergraduate students' beliefs about the random nature of molecular environments. Experiment 1 involved the implementation of a pilot adaptive assessment ($n = 773$) and focus group discussions with undergraduate students enrolled in first- through third-year biology courses; experiment 2 involved the distribution of the redesigned adaptive assessment to the same population of students in three consecutive years ($n = 1170$). The overarching goal of the study was to provide a detailed characterization of learners' perceptions and beliefs regarding molecular agency, environments, and diffusion and whether or not those beliefs change over time. Our results indicated that advanced learners hold as many misconceptions as novice learners and that confidence in their misconceptions increases as they advance through their undergraduate education. In particular, students' understanding of random/Brownian motion is complex and highly contextual, suggesting that the way in which we teach biology does not adequately remediate students' preconceived notions of molecular agency and may actually reinforce them.

INTRODUCTION

It is well established that students in undergraduate biology struggle with understanding a number of complex and abstract scientific concepts, many of which are operating at spatial and temporal scales for which students have no real-world referents. Students bring with them scientifically inaccurate ideas from their prior experience, many of which are firmly established and therefore persistent (Chi, 2005, 2008; Schönborn and Bögeholz, 2009). They are particularly challenged to understand how random mechanisms contribute to the functioning of complex, seemingly efficient, cellular systems. In attempting to reconcile these apparent differences, they will compensate by attaching agency, or directedness, to molecular entities (Chi and Roscoe, 2002; Slotta and Chi, 2006; Chi *et al.*, 2012; Garvin-Doxas and Klymkowsky, 2008; Momsen *et al.*, 2010; Jenkinson and McGill, 2012, 2013). In particular, students experience difficulty comprehending the abstract concept of random or emergent processes taking place within a cellular environment. Rather, they impose their understanding of straightforward, mechanistic systems upon the molecular realm (Chi, 2013). Comprehending the dynamic nature of biological systems and the structures that enable them to interact is critical for a solid understanding of living organisms. Many students will graduate without an understanding of these fundamental concepts (Nehm and Reilly, 2007; Andrews *et al.*, 2011).

While a number of the misconceptions associated with life sciences education have been broadly identified (Ellis, 2001; Garvin-Doxas and Klymkowsky, 2008;

John Coley, *Monitoring Editor*

Submitted Dec 13, 2017; Revised Oct 23, 2018;

Accepted Oct 26, 2018

CBE Life Sci Educ March 1, 2019 18:ar4

DOI:10.1187/cbe.17-12-0267

*Address correspondence to: Jodie Jenkinson (j.jenkinson@utoronto.ca).

© 2019 A. Gauthier *et al.* CBE—Life Sciences Education © 2019 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

Robic, 2010), the challenge of remediating these misconceptions remains. Understanding the beliefs that students hold regarding biological phenomena is essential to confronting their mental models and initiating conceptual change. Thus, there is a pressing need to more closely examine and characterize these learning challenges so that student difficulties can be prevented or remediated and better-quality learning resources can be developed for biology instructors and their students.

Measuring Student Thinking

Assessment of understanding may serve a number of functions. It may be used to collect formative feedback, as a summative assessment on student performance, or as a self-assessment tool to be used by students to test their own ability. In traditional testing scenarios, students demonstrate mastery by selecting correct responses from a preset list of questions. While these methods may test declarative knowledge of broad concepts, they fail to capture nuanced patterns in understanding (Odom, 1995; Klymkowsky and Garvin-Doxas, 2008; Klymkowsky *et al.*, 2010; Shi *et al.*, 2010; Fisher *et al.*, 2012). Particularly in life sciences education, traditional assessments often fail to capture high-quality information about students' mental models of complex dynamic systems. Assessment methods aimed at revealing learners' underlying thought processes and beliefs (e.g., interviews, oral examinations, or demonstrations) are effective for collecting rich qualitative data. However, these methods are outside the skill set of many instructors, and the results are very time-consuming to evaluate (Newman *et al.*, 2016). A recent approach to identifying and categorizing misconceptions has involved the development of concept inventory assessment tools. These instruments address subject matter in a growing number of areas, including general biology (Klymkowsky *et al.*, 2010), genetics (Smith *et al.*, 2008), and meiosis (Kalas *et al.*, 2013). Concept inventories have been very successful diagnostic tools in identifying misconceptions and areas that would benefit from more targeted instruction. However, a more granular assessment of misconceptions held by life sciences students is needed to design successful remediation.

Considering Assessment Alternatives—Computerized Adaptive Assessment

Computerized adaptive testing (CAT) is a form of measurement in which questions are posed to students one at a time. The presentation of each test item is dynamically adapted based upon students' answers (van der Linden and Glas, 2000). The goal of CAT is to optimize the testing process and to measure student ability (Huang *et al.*, 2009). CAT has also been used as a diagnostic tool (Guzmán and Conejo, 2004) and more recently has been proposed as an approach for measuring misconceptions (Guzmán and Conejo, 2015). In this paper, we describe the development and testing of the Molecular Concepts Adaptive Assessment tool, which builds upon our research program examining students' understanding of molecular environments and the development of visualization strategies to support learning in life sciences education.

Overarching Research Aims and Structure

This research aims to characterize students' misconceptions surrounding random molecular motion and interactions, crowded molecular environments, and diffusion by using adaptive

assessment in two experiments. These concepts, which are essential to comprehending a wide range of biological functions, have been identified as problematic (Garvin-Doxas and Klymkowsky, 2008; Rundgren and Tibell, 2009; Tibell and Rundgren, 2010; Couch *et al.*, 2015) but have proved challenging to remediate (Jenkinson and McGill, 2012, 2013). More research is needed to better understand the nature of students' difficulties with this subject matter (Tibell and Rundgren, 2010). Experiment 1 involves the deployment of a pilot version of a custom-developed, computerized adaptive assessment to undergraduate biology students in first-, second-, and third-year courses. To better understand students' responses and how the assessment could be improved to more accurately reflect their understanding, a focus group was held, and the survey was then revised accordingly. Experiment 2 encompasses the deployment of the finalized assessment to the same population across three consecutive academic years, in which we explore misconceptions and the robustness of misconceptions longitudinally across years of survey distribution and between students enrolled in novice through advanced biology courses.

EXPERIMENT 1 (E1)

E1 Aim, Research Questions, and Hypotheses

The aim of this first experiment (E1) was to gauge the accuracy of an adaptive assessment in characterizing students' misconceptions about molecular emergence by comparing survey responses with focus group discussion outcomes. The findings were intended to be used to redevelop the survey for broader distribution in experiment 2 (E2).

We asked the following questions in E1:

- Q1. How did students respond to each MCAA item, and did responses vary between students enrolled in introductory, intermediate, and advanced biology courses?
- Q2. Did responses reflect students' true understanding of the concepts or, rather, a misunderstanding of the question?
- Q3. What changes (e.g., different wording, different adaptive structure, new items, new response options) should be made to the MCAA before a wider distribution in E2?

We hypothesized that a large proportion of students would respond with misconceptions to each MCAA statement and that misconceptions would, in general, be held equally among students enrolled in introductory, intermediate, and advanced biology courses. This view arose from our experience examining misconceptions in molecular biology education (Jenkinson and McGill, 2012, 2013; Gauthier and Jenkinson, 2015, 2017) as well as our interpretation of the literature, which reports that many of these misconceptions are very difficult to resolve (Garvin-Doxas and Klymkowsky, 2008; Rundgren and Tibell, 2009; Tibell and Rundgren, 2010; Couch *et al.*, 2015) and in some cases persist throughout undergraduate education (Nehm and Reilly, 2007; Andrews *et al.*, 2011). While we hoped that the assessment would reflect students' understanding, we expected that our focus group discussions would help us identify ways in which the MCAA items could be improved.

E1 Method

Participants. Participants were undergraduate students from first-, second-, and third-year biology courses at the University of Toronto Mississauga. First-year students, who had little to no

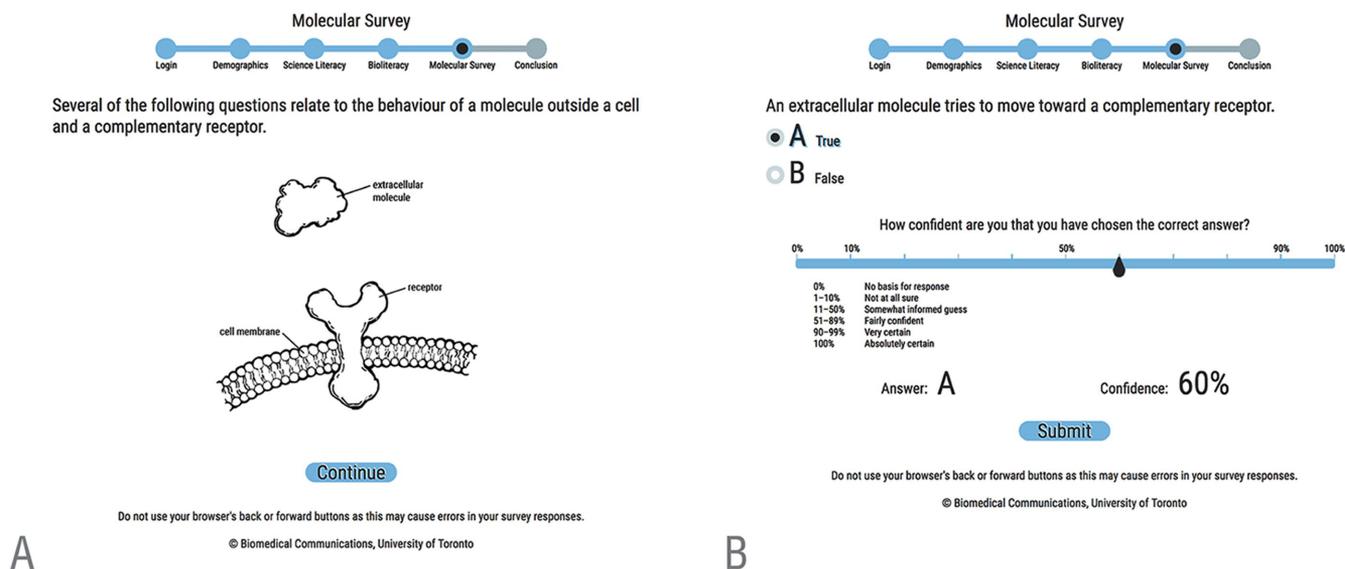


FIGURE 1. GUI of final online study survey (E2). The GUI was similar in E1, apart from the presence of the bio-literacy component of the survey, which can be seen in the progress indicator at the top of the screen. (A) Contextual image before beginning the MCAA portion of the survey; (B) example of presentation of misconception statement with confidence slider.

postsecondary biology background, were enrolled in an introductory biology course that covered concepts regarding evolution, transmission, and evolutionary genetics. Second-year students were enrolled in an intermediate biology course entitled Introductory Cell and Molecular Biology that covered topics including the structure and function of cellular components, the central dogma, protein targeting, genome organization, and more; the concept of Brownian motion was also introduced. Third-year students were enrolled in a more advanced molecular biology course that expanded on topics introduced in previous years, including genetic organization and transmission, DNA replication and repair, and mechanisms of recombination. Each lower-level course was a prerequisite for enrollment in the next higher-level course.

Materials. The study was administered via a password-protected, online survey that students completed remotely. The demographics section of the survey collected information regarding the gender, course of enrollment, major area and year of study, and number of biology and chemistry courses completed by the student. This was followed by a basic science literacy questionnaire with 10 true–false questions (Miller, 1998); this component was included as a baseline measurement of science understanding to control for the possibility that molecular misconceptions might be related to overall competency in science at various levels of undergraduate education. The pilot version of the Molecular Concepts Adaptive Assessment (MCAA), developed jointly by the University of Toronto Mississauga Science Visualization Lab and the Center for Molecular and Cellular Dynamics at Harvard Medical School, was subsequently introduced. An example of the graphical user interface (GUI) is presented in Figure 1. The MCAA portion of the survey began with a contextual image of an extracellular molecule and a complementary receptor (binding partners) as seen in Figure 1A. The survey questions consisted of a maximum of 16 true–false and multiple-choice statements, addressing

14 misconceptions relating to 1) molecular motion, 2) interactions with other molecules within a crowded environment, and 3) diffusion (Figure 1B). The assessment was adaptive, in that some of the statements presented to the students were dependent on their previous answers. For example, if the participant incorrectly answered “True” to the statement “An extracellular molecule attempts to move toward its receptor (i.e., it has a binding objective),” the student was directed to statements that attempt to better characterize the nature of his or her misconception(s). The tree-like structure of the survey can be viewed in Figure 2. For each MCAA statement, the students were asked to rank their confidence in their answers on a continuous slider from 0% (no basis for response) to 100% (absolutely certain). Answers could not be revisited once submitted; the browser’s backward navigation functionality was disabled on the website.

Protocol

MCAA Pilot Distribution. Participants were recruited through a 2-minute presentation made by S.J. at the start of a lecture in the second week of the Fall 2014 semester, and a link to the study website was provided on each course’s BlackBoard portal. Participants logged onto the study website remotely using their student IDs and passwords, accepted the terms delineated on the informed consent form, and proceeded to complete all aspects of the survey. The survey was open for 2 weeks. Responses were stored in a MySQL database on a departmental server, and science literacy and MCAA components were automatically graded using a PHP script. The data for any participant who did not complete the survey in its entirety were discarded. Students who completed the survey were rewarded with a 0.5% bonus mark toward the final biology course grade.

Focus Group. In the week following survey completion, participants were then invited to participate in two focus groups to discuss their experiences and their understanding of the survey questions. The focus groups were facilitated by two investigators

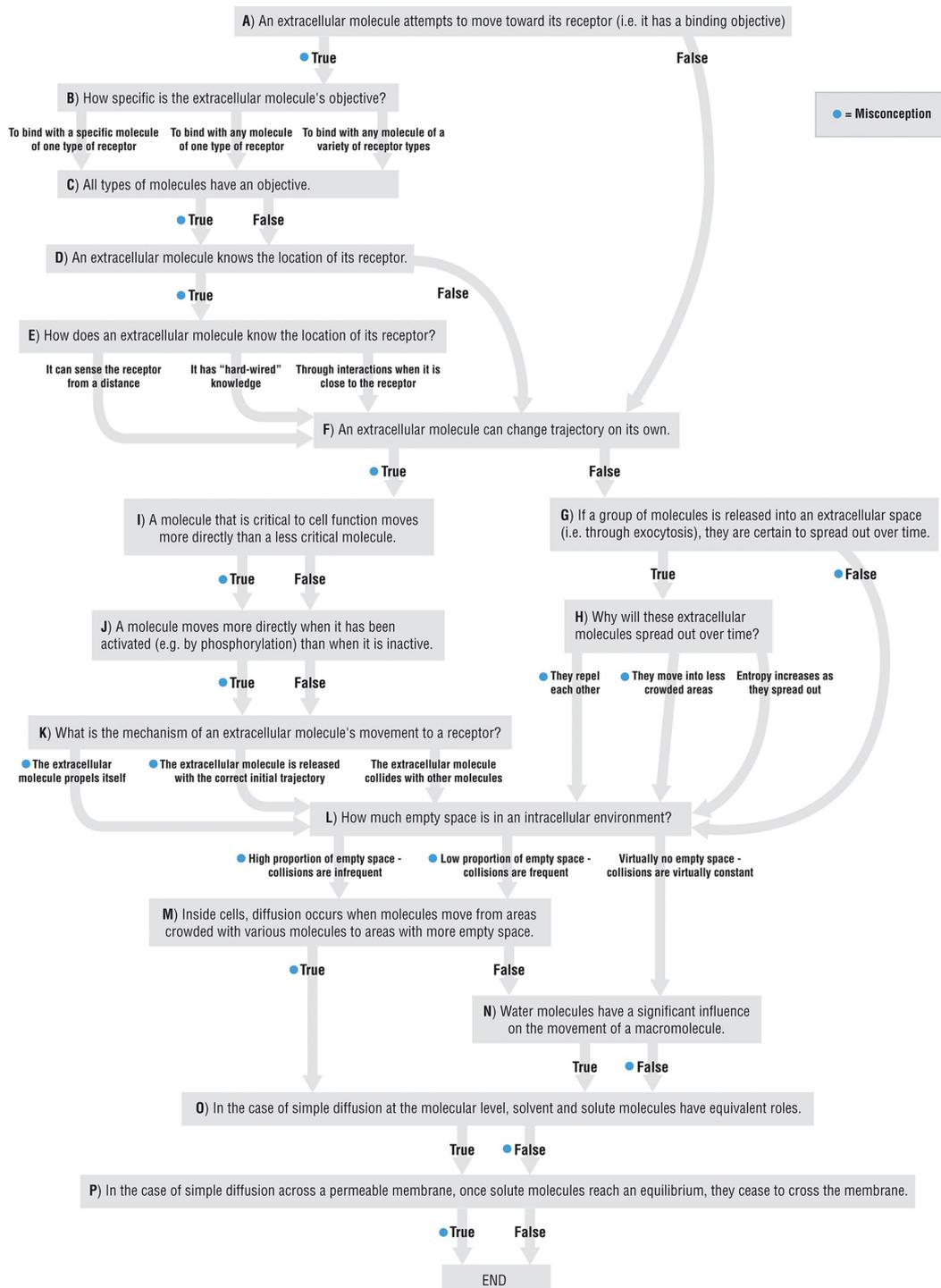


FIGURE 2. Structure of the pilot MCAA used in E1.

who followed a brief discussion guide over 60 minutes (guide provided in the Supplemental Material). The sessions focused on the students' general understanding of the terminology and wording of each question and identification of any additional misconceptions not previously included in the MCAA. Our goals in the focus groups were to determine what changes needed to be made to the wording and/or the adaptive structure of the

survey. The focus group discussions were recorded and transcribed verbatim. Participants were compensated for their time with a \$10 gift card to the university's bookstore.

Data Analysis

General. All statistical analyses in E1 and E2 were performed using SPSS v. 24 (IBM Corporation, 2017) and Excel v. 16.11.

The detailed results of all analyses are reported in the Supplemental Material. When applicable, and across analyses in E1 and E2, effect sizes are provided to characterize the degree of significance of our results, because our sample is large; refer to Supplemental Table 1 for an interpretation guide and rationale for the different effect sizes used.

Effects of Prior Science Knowledge. As a baseline assessment of students' understanding of basic science, differences in science literacy between introductory through advanced biology courses were assessed with a nonparametric Kruskal-Wallis test (the measurements were skewed toward high scores and failed tests of normality). Follow-up pairwise comparisons were performed with Mann-Whitney *U*-tests. The relationship between total misconceptions and science literacy scores was assessed with Spearman correlations.

Q1. How Did Students Respond to Each MCAA Item, and Did Responses Vary between Students Enrolled in Introductory, Intermediate, and Advanced Biology Courses? The results for each MCAA statement were analyzed with a Pearson chi-square test performed to compare frequency of misconceptions between biology courses. We used the Holm's sequential Bonferroni procedure (Holm, 1979) to calculate an alpha value ($HB-\alpha$) for each of the 16 statements (Supplemental Table 2). Therefore, only a *p* value falling below the $HB-\alpha$ for that statement would be considered significant. Post hoc pairwise comparisons were performed on significant results with a new $HB-\alpha$ based on that statement's original ranking. The distribution of responses in multiple-choice characterization questions (statements B, E, H, K, and L) were investigated by chi-square goodness-of-fit tests (using $HB-\alpha$ values).

Q2. Did the Above Responses Reflect Students' True Understanding of the Concepts or, Rather, a Misunderstanding of the Question? In alignment with our second research question, the focus group transcripts were coded for four themes: 1) misinterpretations of the language or meaning of MCAA survey items, 2) evidence for misconceptions identified by the MCAA, 3) evidence for additional misconceptions (not addressed in the MCAA), and 4) demonstration of correct understanding of concepts. General discussion that did not fit into these four themes was coded with a "0." Excel v. 16.11 was used to parse each participant response or single thought into its own row, and two reviewers independently assigned one of the four codes (or "0") to each row in a fully crossed design. Statements made by the focus group facilitators were not coded. In SPSS, kappa's coefficient was calculated to determine interrater reliability, with a value of 0.70 or greater deemed acceptable. Disagreements between reviewers were then resolved until 100% agreement was reached.

Q3. What Changes (e.g., Different Wording, Different Adaptive Structure, New Items, New Response Options) Should Be Made to the MCAA before a Wider Distribution in E2? The results from Q2 were used to inform the restructuring and/or rewording of MCAA items.

E1 Results and Discussion

Results of the pilot assessment, in particular focus group interviews, helped to advance the design of the MCAA by providing

TABLE 1. Science literacy scores (maximum score = 10) and total misconceptions (as measured by the MCAA, maximum misconceptions = 14) in E1, stratified by biology course level

	<i>n</i>	Minimum, maximum	Mean	SD
A. Science literacy				
First year	424	3, 10	8.67	1.29
Second year	277	5, 10	9.05	1.10
Third year	72	6, 10	9.17	1.01
B. Total misconceptions				
First year	424	0, 14	7.42	2.77
Second year	277	0, 14	7.10	2.92
Third year	72	0, 13	5.81	2.75

additional insights into students' interpretation of questions. The survey was revised to clarify and provide greater context for language that was poorly understood. Focus group interviews were also helpful in confirming misconceptions identified in responses to the survey instrument, validating the appropriateness of many questions. Finally, the focus group discussion was instrumental in identifying additional misconceptions not previously included in the MCAA. The findings and implications of the pilot assessment are discussed in *MCAA Pilot Distribution*.

MCAA Pilot Distribution

Participant Composition. A total of 773 students completed the survey at the beginning of the semester, with 424 first-year students, 277 second-year students, and 72 third-year students. Of these participants, 67.3% were female, 32.2% were male, and 0.5% did not disclose their gender (a 3:2 ratio of female:male enrollment has been found to be typical in undergraduate biology classrooms; Eddy et al., 2014).

Effects of Prior Science Knowledge. Participants had an overall high degree of science literacy (scores displayed in Table 1A). While second- and third-year students had higher science literacy than first-year students, the effects were very small (Supplemental Tables 3 and 4). The distribution of misconceptions held by first- through third-year students is displayed in Table 1B. A weak negative relationship existed between total misconceptions and science literacy in first-year students ($r_{sp} = -0.12$, $p = 0.015$) and second-year students ($r_{sp} = -0.15$, $p = 0.013$), but not in third-year students ($r_{sp} = -0.16$, $p = 0.191$). This suggests that misconceptions may ameliorate with an increased understanding of general science, but not to any great extent.

Q1. How Did Students Respond to Each MCAA Item, and Did Responses Vary between Students Enrolled in Introductory, Intermediate, and Advanced Biology Courses? Of 14 misconceptions included in the MCAA, 11 were confirmed as misconceptions by more than 50% of participants. Specific values of participant misconceptions for each MCAA pilot statement can be viewed in Table 2. The detailed results of chi-square tests for each misconception statement are provided in Supplemental Table 5, with post hoc results in Supplemental Table 6. Results for any multiple-choice characterization questions (B, E, H, K, and L) are provided in Supplemental Table 7, with post hoc results in Supplemental Table 8.

TABLE 2. Distribution of correct responses (✓) and misconceptions (✗) on MCAA survey in E1

Pilot statements (E1: 2014)—Correct answer in bold		Total n	✓ n (%)	✗ n (%)
A	An extracellular molecule attempts to move toward its receptor (i.e., it has a binding objective). (True or False)	773	130 (16.8)	643 (83.2)
B	How specific is the extracellular molecule's objective?	—	—	—
	I. To bind with a specific molecule of one type of receptor ($n = 444$, 69.1%)			
	II. To bind with any molecule of one type of receptor ($n = 167$, 26.0%)			
	III. To bind with any molecule of a variety of receptor types ($n = 32$, 5.0%)			
C	All types of molecules have an objective. (True or False)	643	218 (33.9)	425 (66.1)
D	An extracellular molecule knows the location of its receptor. (True or False)	643	257 (40.0)	386 (60.0)
E	How does an extracellular molecule know the location of its receptor?	—	—	—
	I. It can sense the receptor from a distance. ($n = 37$, 9.6%)			
	II. It has “hard-wired” knowledge. ($n = 63$, 16.3%)			
	III. Through interactions when it is close to the receptor. ($n = 286$, 74.1%)			
F	An extracellular molecule can change trajectory on its own. (True or False)	773	525 (67.9)	248 (32.1)
G	If a group of molecules is released into an extracellular space (i.e., through exocytosis), they are certain to spread out over time. (True or False)	525	433 (82.5)	92 (17.5)
H	Why will these extracellular molecules spread out over time?	433	200 (46.2)	233 (53.8)
	I. They repel each other. ($n = 46$, 10.6%)			
	II. They move into less crowded areas. ($n = 187$, 43.2%)			
	III. Entropy increases as they spread out. ($n = 200$, 46.2%)			
I	A molecule that is critical to cell function moves more directly than a less critical molecule. (True or False)	248	114 (46.0)	134 (54.0)
J	A molecule moves more directly when it has been activated (e.g., by phosphorylation) than when it is inactive. (True or False)	248	50 (20.2)	198 (79.8)
K	What is the mechanism of an extracellular molecule's movement to a receptor?	248	107 (43.1)	141 (56.9)
	I. The extracellular molecule propels itself. ($n = 64$, 25.8%)			
	II. The extracellular molecule is released with the correct initial trajectory. ($n = 77$, 31.0%)			
	III. The extracellular molecule collides with other molecules. ($n = 107$, 43.1%)			
L	How much empty space is in an intracellular environment?	773	169 (21.9)	604 (78.1)
	I. High proportion of empty space—collisions are infrequent. ($n = 177$, 22.9%)			
	II. Low proportion of empty space—collisions are frequent. ($n = 427$, 55.2%)			
	III. Virtually no empty space—collisions are virtually constant. ($n = 169$, 21.9%)			
M	Inside cells, diffusion occurs when molecules move from areas crowded with various molecules to areas with more empty space. (True or False)	604	98 (16.2)	506 (83.8)
N	Water molecules have a significant influence on the movement of a macromolecule. (True or False)	267	220 (82.4)	47 (17.6)
O	In the case of simple diffusion at the molecular level, solvent and solute molecules have equivalent roles. (True or False)	773	384 (49.7)	389 (50.3)
P	In the case of simple diffusion across a permeable membrane, once solute molecules reach an equilibrium, they cease to cross the membrane. (True or False)	773	383 (49.5)	390 (50.5)

Focus Groups

Participant Composition. The focus groups included eight students ($n = 5$ first-years, $n = 1$ second year, and $n = 2$ third years) over two sessions ($n = 4$ in each). The first focus group consisted of Participants 1–4, and the second focus group of Participants 5–8.

Codes and Interrater Reliability. In total, 167 participant statements were parsed and coded. A full list of extracted coded statements from the focus group transcripts are provided in Supplemental Table 34. Interrater reliability was considered excellent, with 88% agreement and $\kappa = 0.835$, 95% confidence interval [0.767, 0.903]. Focus group outcomes are organized by code category.

Q2. Did the Above Responses Reflect Students' True Understanding of the Concepts or, Rather, a Misunderstanding of the Question?

Evidence of the misinterpretation of MCAA statements and response options. The wording of several MCAA items posed a challenge for participants. For statement A, it appeared that the presence of the phrase “it has a binding objective” acted as a primer for students and may have erroneously influenced their responses. Participant 8, a second-year student, stated, “I feel in class ... we usually learn about how different molecules have a specific job in the cell, so automatically I think: it does have a specific objective” and that “from what I've learned so far, everything that we've talked about has some kind of objective; maybe it isn't carrying out that objective right away or in that moment, but they have something that they tend to do

within a cell.” It was generally agreed among participants that the wording to the follow-up statement B (refer to Figure 2) was confusing; they did not understand that multiple copies of identical receptors could be present. This is well demonstrated by Participant 1 (third-year), who stated, “To bind any molecule of one type of receptor, uh, to me I sort of thought of it as: you have a type of receptor [...] which is made up of different molecules, and your particular extracellular molecule could bind to any of those different types, as opposed to [...] one type of receptor that’s identical and you can bind to any of them.”

Statement F, regarding a cell’s ability to prioritize “critical” processes over others, was misleading for some students. While Participant 1 interpreted the statement as we intended, reasoning “Cause... if [the molecule] needs to do something, then the cell would have a reason to make sure it completes that function as opposed to something that is less critical,” Participant 8 identified the question as confusing, asking, “What you would consider less critical as compared to critical? [...] Because in my head, I would think they’re all kind of critical, somehow.”

The concept of “empty space” in the cell also posed interpretive challenges (statement L). While participants generally responded that there was a low proportion of empty space in cells (rather than no empty space at all), it appeared that their use of the term “empty space” referred to regions with low concentrations of macromolecules but where water and other small molecules may still be present. For example, Participant 8 said, “I don’t know if by empty space do you mean space that there’s absolutely nothing, or can there be some kind of like, say like cytoplasm [...] I know [the cytoplasm]’s a fluid, but ... I would consider this empty space, but really it’s not, right?” Similarly, for statement Q, students expressed concern about our use of terminology when describing the “roles” of solvent and solute molecules: “I’m kind of confused about the ‘equivalent roles’; like, I know each of them have a role, but [...] what do you mean by that?” (Participant 7, third-year).

Evidence of misconceptions identified by MCAA items.

Throughout the focus group discussions, students’ responses revealed a surprisingly high level of misunderstanding regarding molecular agency.

This was exemplified when students were asked about statement D: “An extracellular molecule knows the location of its receptor.” Participant 3 (first-year), when discussing the statement in the context of DNA replication, suggests, “We need to have all the complexes bind to the DNA strand so that it can start the DNA replication process ... you have all these complexes coming from different parts of the cell, and they obviously must know where the DNA strand is in order to come together in such a way that it can function in DNA replication, so I would say yes [to the statement].” The students were questioned further about how they thought molecules communicated with each other, and the most prevalent theme was long-distance forces of attraction: Participant 4 (first-year) said, “[In the context of blood vessels] they would be brought through the blood and then if they came into proximity of the receptor then they would have the attraction and they would bond there,” while Participant 2 (first-year) reasoned that “They’re always just going to attach to that certain receptor regardless of distance.”

The focus groups also highlighted that students’ understanding of Brownian motion was highly contextual. Some students

applied concepts of random motion only to scenarios wherein molecules were not concurrently interacting in cellular processes; for example, Participant 2 stated, “Randomized, I feel like it doesn’t have a purpose, so it’s not moving in order to establish a bond or to ... accomplish something? I just think it’s floating, and when it’s directed, I feel like either the molecules are binding to do something, or they have a purpose to their movement, so... they’re trying to accomplish something.” Extending this theme, when discussing statement I (activation leads to more direct motion), participants described how the floating behavior changes upon phosphorylation. For instance, Participant 7 stated, “When a molecule is activated, like it does go to where it needs to go; when it’s inactivated, it will just sit there and not really have a purpose.”

While individuals in the second focus group expressed confusion regarding statement Q (as described in the preceding section), those in the first focus group displayed a mixed understanding of diffusion and the roles of solute and solvent molecules. While many understood the equivalency of molecular behavior in diffusion (e.g., Participant 2 said, “I feel like they’re pretty much equivalent ‘cause both of them are contributing the same, like at a molecular level, they’re both interacting with one another”), others (e.g., Participant 3) believed that only the solute molecules diffused: “linking back to like the basic definition of diffusion, so from particles moving from higher to lower concentration, I would say that the solute has a bit more of an important role in the sense where it’s the one moving and not the solvent itself?”

Evidence of misconceptions not already identified by MCAA items.

A striking new misconception identified through the focus groups is that molecules receive messages from the nucleus to direct or initiate their required function in the cell. Participant 5 (a first-year) said, “I think there’s specific messengers; we know the core of the cell is the nucleus, which is able to carry out the messages and to convey whether *this* needs to join with *this* to perform *this* function, so I think everything starts from the nucleus” and that “sometimes some molecules already have the function ingrained in themselves; they’re able to carry out [the function] on their own, they just require that message from the nucleus to kind of do the function.” Participant 6 (first-year) agreed, saying, “Because the way we were taught, the nucleus was kind of like the brain of the cell, where it tells the molecules what to do ... it directs the molecules to do everything, so the molecules themselves wouldn’t do anything, unless the nucleus specifically sends out a message.”

A second misconception not identified by the original version of the MCAA was in regard to statement K (the mechanism of molecular movement toward a binding partner); few students reasoned that a molecule’s primary mechanism of motion is through random collisions. A prominent line of reasoning was that molecules were carried by other “helper” molecules or that “something else brings it by” (Participant 4). In the context of protein translation, Participant 1 said “the tRNA would bring amino acids to [the] ribosome so, in that sense, the amino acid’s being directed to the ribosome by tRNA for the overall purpose of synthesizing a protein.” Participant 3 suggested that “everything has [...] a purpose to do something, so I guess it kind of depends on each other in a sense; so, one thing needs to be made before the other can be made and then those two can

maybe like help each other kind of get to where they need to go, so I guess it's really dependent on one another, so that's not so much independent work but more dependent."

Finally, it became apparent that, while students applied teleological reasoning to larger molecules like proteins, these misconceptions did not necessarily extend to smaller molecules: "But then again, it's also weird to think about like, when we're talking about DNA and proteins, but then when you talk about just [...] inorganic molecules such as carbon dioxide or other things ... that's a little bit trickier to understand" (Participant 1).

Evidence of correct understanding of concepts. Students generally had a good understanding that the structure of a molecule mediates its function (e.g., lock-and-key mechanisms). Interestingly, none of the focus group participants exhibited any confusion about the continuity of molecular motion during diffusion, though a large proportion of survey respondents answered incorrectly to statement R. Furthermore, some participants had a better understanding of role of random mechanisms than others; in particular, Participant 8 showed a good understanding that activation did not mediate the nature of movement and rate of collisions with binding partners when both activated and nonactivated versions were present: "I wanna say, [...] like they would be equally in contact, it was just that the one that is actually active would actually connect, because [...] I guess it depends on the amount of [...] molecules that are there" and "there is really no [...] direct movement? I think they're all just kind of random collisions in a way."

Q3. What Changes (e.g., Different Wording, Different Adaptive Structure, New Items, New Response Options) Should Be Made to the MCAA before a Wider Distribution in E2? The focus groups highlighted the need for restructuring of the assessment and rewording some of the statements. The wording of each pilot statement can be viewed beside each new statement in Supplemental Table 9. The following changes were made:

- The term "binding objective" was interpreted by students to mean that a molecule had a function in the cell, which prompted students to answer with a misconception, regardless of whether they could identify random collisions as the mechanism of molecular motion. As such, the word "objective" was removed from statements A and B, while statement C was removed altogether.
- The answer options in statement B were reworded for clarity.
- We agreed with Participant 8 that the concept of "critical" processes is, in itself, a flawed concept; therefore, statement F was replaced with a new but related statement based on how students associate random motion with small molecules (like water and carbon dioxide) more so than with larger molecules (like proteins).
- Because students refer to the cell's cytoplasm as "empty," even though they understand that it is filled with water and other molecules, the concept of "empty space" (statement L) was better articulated to eliminate possible misinterpretation.
- The "roles" or behaviors of solute and solvent molecules (statement Q) were also rearticulated so that students could make a direct comparison between them.
- The concept that the nucleus could communicate with and direct molecules was included in the redesign as an answer option to statement D, which attempts to characterize

students' misconception that molecules know the locations of complementary binding partners.

- The concept that molecules require help to get to a binding partner was added as an answer option to statement K.
- Statement K (regarding the mechanism of molecular motion) was found to be a core misconception and would be moved higher up in the structure of the survey, so that it was presented to all participants; other statements (e.g., statements F, I, and J) would follow K to help better characterize the misconception.
- Statement G was modified to more directly follow up on students' understanding of random collisions (statement E); statement H was excluded, as it did not fit with the changes made to G.
- Statement N was excluded, because its concept was redundant with statements K and M.
- Finally, several of the questions were slightly reworded to ensure that there was no ambiguity in their meaning (e.g., the term "trajectory" was replaced with "direction" in statement F).

The structure of the new survey as used in Experiment 2 is visualized in Figure 3.

EXPERIMENT 2 (E2)

E2 Aim, Research Questions, and Hypotheses

The aim of Experiment 2 (E2) was to characterize students' misconceptions—and their certainty in those misconceptions—across a large sample and to make comparisons between educational levels, as well as longitudinally over a single semester and over three consecutive years.

We asked the following research questions in E2:

- Q1. How does the overall frequency of misconceptions held by students change over time?
- Q2. How did students respond to each MCAA item, and which misconceptions change over time?
- Q3. How strongly do students adhere to their misconceptions, and does their confidence change over time?
- Q4. How did the rewording/restructuring of the MCAA influence students' responses?

As with E1, we hypothesized that the frequency of misconceptions regarding molecular emergent systems would remain constant over time, that students would be confident in their misconceptions, and that this confidence might increase with exposure to more advanced biology courses.

Furthermore, we anticipated that the responses to MCAA items would be different in comparison with matched E1 pilot statements in which intentional changes to the wording/structure of the assessment were implemented.

E2 Methods

Participants. New participants were drawn from the same first-, second-, and third-year biology courses as in E1.

Materials. The demographics and science literacy components of the survey were the same as those used in E1. Because science literacy scores in E1 were very high, the new survey also included a 10-question, multiple-choice, bio-literacy component to help us understand more specifically students' prior biology knowledge and whether it is related to misconceptions. Whereas the

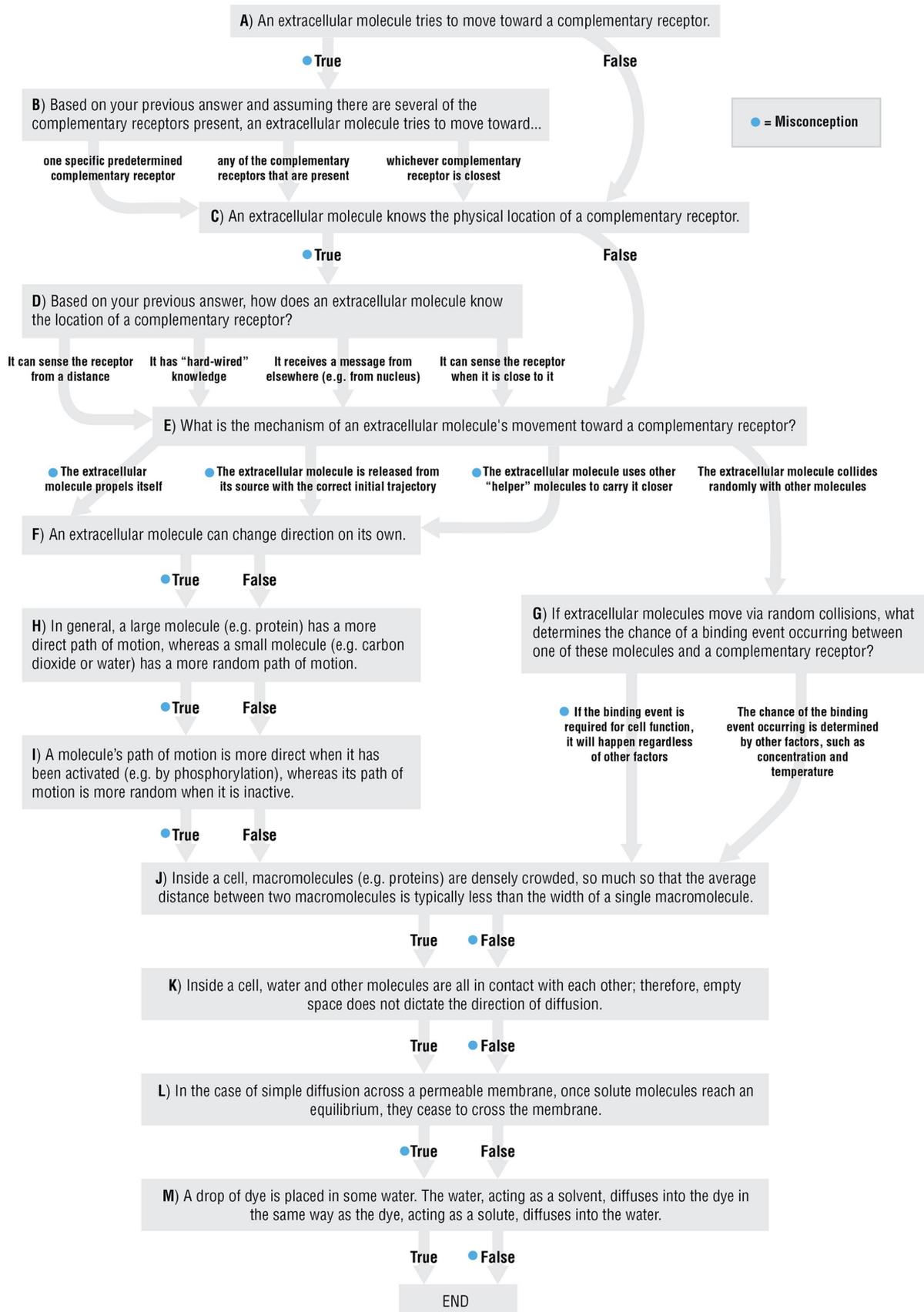


FIGURE 3. Structure of the final MCAA used in E2.

science literacy measure assessed general knowledge, the bio-literacy survey measured knowledge more closely related to cell biology. Questions addressed students' understanding of energetics and interactions, natural selection, genetic drift, and molecular properties. Questions were drawn from Biology Concepts Inventory (Klymkowsky *et al.*, 2010), but questions regarding molecular dynamics were excluded to avoid redundancy with and priming for the MCAA. The finalized MCAA survey (Figure 3) consisted of a maximum of 11 misconceptions over 13 possible statements.

Protocol. The revised assessment was distributed in the 2015 and 2016 academic years, at the beginning and end of the semester, following the same protocol as described for E1. Any E1 participants who may have been enrolled in our targeted courses during E2 were allowed to complete the survey for the bonus mark incentive but were excluded from E2 analyses. E2 participants could complete the survey in both distribution years; for between-subjects analyses that combine 2015 and 2016 data, only participants' first responses (i.e., 2015 survey and not 2016 survey) were used to avoid unit of analysis errors. In the 2017 academic year, we followed up with E2 participants who participated in both 2015 and 2016 with another distribution of the survey (new participants were not recruited; those who were emailed for follow-up were not necessarily enrolled in one of our original target courses any longer). Those who completed the follow-up survey were incentivized with a chance to win one of twenty \$10 gift cards to an on-campus coffee shop through a random draw. Participants with missing data were excluded list-wise from repeated-measures analyses.

Data Analysis

Effect of Prior Science Knowledge. While not a primary research question in this experiment, participants' understanding of basic science (science literacy) and biology (bio-literacy) were measured to confirm that misconceptions are not present due to the absence of general knowledge. In both science and bio-literacy scores, Kruskal-Wallis tests were used to compare prior knowledge across educational levels, while Spearman correlations were used to assess their linear relationship with total misconceptions.

Furthermore, a comparison of science literacy from E1 to E2 can help to assure the similarity of participants across both experiments; this was assessed with a Mann-Whitney *U*-test at each educational level (i.e., first, second, and third year).

Q1. How Does the Overall Frequency of Misconceptions Held by Students Change over Time? Change in misconceptions over time was assessed in three ways: 1) between introductory, intermediate, and advanced biology courses; 2) over the course of a single semester; and 3) over three consecutive academic years (i.e., pretests in 2015, 2016, 2017) with repeated-measures participants.

Between introductory, intermediate, and advanced courses. We performed a univariate analysis of variance (ANOVA) to determine whether students enrolled in increasingly advanced courses (first-, second-, or third-year biology) hold fewer misconceptions, while controlling for year of distribution (2015 or 2016) to ensure that our results were consis-

tent in both distribution years. This analysis was performed using pretest data to achieve the largest sample and because pretest data held a normal distribution.

Over the course of a semester. A two-related-samples Wilcoxon test was performed to determine whether the frequency of misconceptions changed over the course of a semester as a result of students' enrolled biology courses (this nonparametric test was performed because a pre- to postsemester repeated-measures ANOVA failed Box's test of equality of covariance matrices [Box's $M = 27.94$, $F(15, 6107.44) = 1.79$, $p = 0.030$]). A follow-up Kruskal-Wallis test on pre- to postsemester change (i.e., posttest minus pretest) was performed using educational level as the independent variable.

Because some attrition occurred from pretest to posttest, a simple univariate ANOVA was performed to determine whether the reduced sample was representative of the whole group, using total pretest misconceptions as the dependent variable and posttest completion status (yes/no) as the independent variable.

Over three consecutive academic years. We ran a repeated-measures analysis with pretest misconceptions scores in 2015, 2016, and 2017 with individuals who completed the survey in all three academic years to assess how misconceptions change over the academic careers of individuals. Again, to determine whether the reduced sample of individuals who persisted in the study through 3 years was representative of the whole, we performed a simple univariate ANOVA, using total pretest misconceptions (in 2015) as the dependent variable and retained status (yes/no) as the independent variable.

Q2. How Did Students Respond to Each MCAA Item, and Which Misconceptions Changed over Time? Change in each individual misconception was assessed 1) by comparing the distribution of the misconceptions between students enrolled in introductory through advanced courses (similar to E1) and 2) by comparing change in individual misconceptions over the course of a semester.

Between introductory, intermediate, and advanced courses. To help characterize the misconceptions held by the students, we performed the same analyses as described for E1. Again, $HB-\alpha$ values were used to assess significance of the multiple comparisons (Supplemental Table 10).

Over the course of a semester. To better understand which misconceptions were likely to resolve over the course of a semester as a result of typical instruction, we performed McNemar chi-square tests on each individual MCAA statement (correct vs. incorrect answers), using the Holm-Bonferroni method to assess significance.

Q3. How Strongly Do Students Adhere to Their Misconceptions and Does Their Confidence Change over Time? In addition to the change in misconceptions over time (Q1 and Q2), participants' adherence to their misconceptions was assessed through their reported confidence levels in their answers on MCAA items. The data on confidence ratings failed to approximate a normal distribution, so nonparametric tests were used. Confidence in

responses was assessed overall (i.e., average across all MCAA items) and for each MCAA item individually. Change in confidence in misconceptions was assessed 1) between introductory, intermediate, and advanced biology courses (using confidence in correct answers for comparability) and 2) over the course of a semester.

Between introductory, intermediate, and advanced courses.

A Kruskal-Wallis test, using biology course level as the independent variable, was performed on average baseline confidence for 1) incorrect answers (i.e., misconceptions) and 2) correct answers. Furthermore, we performed multiple Kruskal-Wallis tests to investigate differences in confidence between educational levels for each MCAA item, for both correct and incorrect answers separately. Post hoc analyses were performed using Mann-Whitney *U*-tests with HB adjustments. Finally, a within-subjects comparison of confidence in correct versus incorrect answers on MCAA items was made using a Wilcoxon test at each educational level.

Over the course of a semester. To determine how the biology courses themselves might impact students' adherence to their misconceptions, we performed a Wilcoxon assessment on the average confidence reported on *incorrect* answers from pretest to posttest at each educational level. Additionally, we ran the same assessment for confidence on *correct* answers as a means of comparison. Furthermore, we ran individual Wilcoxon tests on each of the 11 misconception statements at each educational level (HB- α based on 33 comparisons total), including only students who *retained* the misconceptions from the start to the end of the semester.

Q4. How Did the Rewording/Restructuring of the MCAA Influence Students' Responses? To determine whether the rewording and restructuring of the MCAA better reflect the focus

group discussions (which we believe is evidence of students' true understanding), we performed Pearson chi-square tests for each of the 11 misconception items from E2 paired with its closest match from E1. Similar analyses could not be performed on multiple-choice characterization items B, D, and E, because these items in E1 and E2 held a different number of response options; however, the goodness-of-fit tests reported in Supplemental Tables 7 (E1) and 22 (E2) were qualitatively compared to come to an interpretation of how changes to the MCAA influenced students' responses to these multiple-choice items.

E2 Results

Participant Composition. In 2015, a total of 506 students participated, with 363 first-year students, 124 second-year students, and 19 third-year students. In 2016, an additional 664 students registered, with 506 first-years, 127 second-years, and 31 third-years—giving us a total sample of 1170 unique participants. Overall, 70.8% of participants were female, 29.0% were male, and 0.2% did not disclose their gender.

At the end of the semester (posttest), 375 (74.11%) participants were retained in 2015, and 506 (76.20%) participants were retained in 2016. This number excludes individuals who, in 2015 ($n = 40$) and 2016 ($n = 84$), participated in other molecular concepts intervention studies and who were not included in repeated-measures analyses here for that reason (Gauthier and Jenkinson, 2015, 2017). The distribution of participants across years at both the pretest and posttest can be viewed in Figure 4. Some students repeated the survey in multiple academic years; 140 first-years and 17 second-years from 2015 also participated in 2016 (now registered in their second and third years, respectively); attrition here is likely due primarily to the students not being enrolled in one of the targeted courses at the time of the study. A small sample of 28 students completed the survey in all three

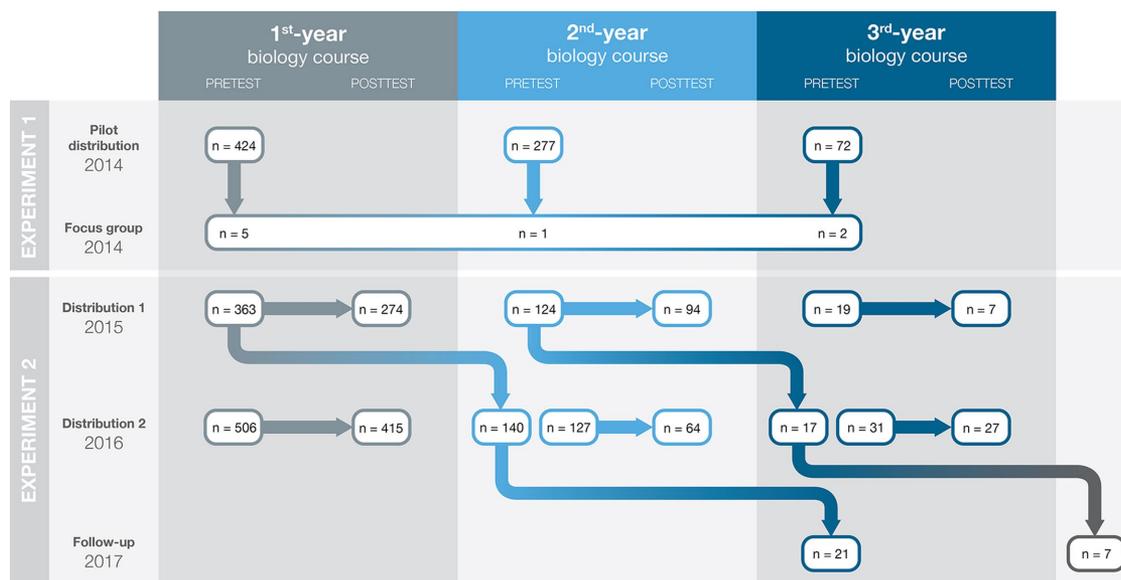


FIGURE 4. Study structure and participant numbers. Two groups are represented at the pretest in 2016 for the second- and third-year biology courses: a group of students who previously participated in 2015 and a group of new participants who had not previously completed the survey. These groups are analyzed separately to avoid unit of analysis errors.

TABLE 3. Science literacy scores (maximum score = 10), bio-literacy scores (maximum score = 10), and total misconceptions (as measured by the MCAA, max misconceptions = 11) in E2, stratified by biology course level

	<i>n</i>	Minimum, maximum	Mean	SD
A. Baseline science literacy				
First year	869	3, 10	8.58	1.30
Second year	251	3, 10	8.79	1.35
Third year	50	5, 10	8.93	1.26
B. Baseline bio-literacy				
First year	869	0, 10	4.60	1.80
Second year	251	1, 10	5.06	1.70
Third year	50	1, 9	5.90	1.85
C. Pretest misconceptions				
First year	869	0, 11	5.96	2.24
Second year	251	1, 11	6.03	2.19
Third year	50	0, 10	5.19	2.74

consecutive academic years. Of these, 21 started off as first-years and seven started as second-years; therefore, in 2017, they would have been in their third and fourth years of undergraduate studies respectively. Attrition in 2017 is likely to be due to the fact that their participation was no longer linked to a course, and therefore they were not eligible for a bonus mark for participation.

Effects of Prior Science Knowledge. Science literacy scores at baseline were again very high (Table 3A) and did not differ from those seen in E1 (Supplemental Table 11). Some differences between educational levels existed (second-year students scored significantly higher than first-year students), but these differences were negligible ($\eta^2 < 0.01$; Supplemental Table 13). Bio-literacy was overall lower than science literacy (Table 3B). Significant differences were observed between educational levels, with each level scoring higher than the level below it (Supplemental Tables 12 and 13).

Science literacy scores were negatively but very weakly correlated with misconceptions in first-year ($r_{sp} = -0.09$, $p = 0.006$) and second-year ($r_{sp} = -0.17$, $p = 0.006$) students, and moderately correlated in third-year students ($r_{sp} = -0.45$, $p = 0.001$). Similarly, bio-literacy scores and misconceptions were negatively but weakly correlated in first-years ($r_{sp} = -0.10$, $p = 0.003$), second-years ($r_{sp} = -0.14$, $p = 0.024$), and third-years ($r_{sp} = -0.31$, $p = 0.030$).

Q1. How Does the Overall Frequency of Misconceptions Held by Students Change over Time?

Between introductory, intermediate, and advanced courses. The distribution of misconceptions at the beginning of the semester (pretest) can be seen in Table 3C (a per item breakdown is shown in Table 4). Results indicate that students at all three levels in undergraduate biology hold a similar number of misconceptions and that results were similar between distribution years (Supplemental Table 14).

Over the course of a semester. Misconceptions measured at the end of the semester (as well as pretest misconceptions for the

reduced sample) are presented in Table 5. An overall significant but very small drop in misconceptions was observed over the course of the semester ($Z = -3.97$, $p < 0.001$, $r_{cc} = -0.13$), with no difference in conceptual change between educational levels ($\chi^2(2) = 0.028$, $p = 0.986$). Though attrition occurred from pretest to posttest, pretest misconception scores were similar between those who did and those who did not complete the posttest (Supplemental Tables 15 and 16), suggesting that the smaller repeated-measures sample is representative of the larger group.

Over the course of three consecutive years. Table 6 displays the total misconceptions of the 28 individuals who completed the MCAA in the 2015, 2016, and 2017 academic years. Our model showed no significant change in misconceptions over three academic years, regardless of whether students started in their first or second year of postsecondary education (Supplemental Table 17). No linear ($F(1) = 1.52$, $p = 0.229$) or quadratic ($F(1) = 0.73$, $p = 0.400$) effects were observed for change over time, nor were linear ($F(1) = 0.36$, $p = 0.555$) or quadratic ($F(1) = 0.40$, $p = 0.531$) effects observed for the interaction between change and educational level. Individuals who participated in all 3 years performed no differently than their classmates who chose to cease participation in 2015 (Supplemental Tables 18 and 19), which suggests that our retained sample is reflective of the whole.

Q2. How Did Students Respond to Each MCAA Item, and Which Misconceptions Change over Time?

Between introductory, intermediate, and advanced courses. Results per MCAA item as measured at the pretest are presented in Table 4, with details for each chi-square test presented in Supplemental Tables 20–23. Generally speaking, rates of misconceptions were high, with greater than 40% of the sample responding with misconceptions to nine out of 11 statements. Misconceptions were equally distributed across first-, second-, and third-year students in all but one statement (C: “An extracellular molecule knows the physical location of a complementary receptor”), which was less prevalent in third-year students.

Over the course of a semester. When assessing change in each misconception from pretest to posttest, two specific misconceptions saw significant improvement, though the effects are considered to be very small to small. First, misconception F, regarding the ability of molecules to change direction, saw the most improvement, with a 46% increased chance of being answered correctly at the end of the semester (OR = 1.464). Second, there was a 30% increased chance that students answered misconception L (about diffusion and equilibrium) correctly on the posttest as compared with the pretest (OR = 1.297). The detailed results for all misconceptions are presented in Supplemental Table 24.

Q3. How Strongly Do Students Adhere to Their Misconceptions and Does Their Confidence Change over Time?

Between introductory, intermediate, and advanced courses. Confidence ratings of 0% indicate that the student made a random guess, while responses greater than 50% indicate that the student was reasonably confident, and 100% indicates that the student was absolutely certain. Overall, participants’

TABLE 4. Distribution of correct responses (✓) and misconceptions (✗) on MCAA survey in E2

ID	Redesigned statements (E2: 2015 and 2016)—Correct answer in bold	Total n	✓ n (%)	✗ n (%)
A	An extracellular molecule tries to move toward a complementary receptor. (True or False)	1170	245 (20.9)	925 (79.1)
B	Based on your previous answer, and assuming there are several of the complementary receptors present, an extracellular molecule tries to move toward: I. one specific predetermined complementary receptor; (n = 524, 56.6%) II. any of the complementary receptors that are present; (n = 184, 19.9%) III. whichever complementary receptor is closest. (n = 217, 23.5%)	925	—	—
C	An extracellular molecule knows the physical location of a complementary receptor. (True or False)	1170	626 (53.3)	544 (46.5)
D	Based on your previous answer, how does an extracellular molecule know the location of a complementary receptor? I. It can sense the receptor from a distance. (n = 46, 8.5%) II. It has “hard-wired” knowledge. (n = 55, 10.1%) III. It receives a message from elsewhere (e.g., from nucleus). (n = 304, 55.9%) IV. It can sense the receptor when it is close to it. (n = 139, 25.6%)	544	—	—
E	What is the mechanism of an extracellular molecule’s movement toward a complementary receptor? I. The extracellular molecule propels itself. (n = 158, 13.5%) II. The extracellular molecule is released from its source with the correct initial trajectory. (n = 192, 16.4%) III. The extracellular molecule uses other “helper” molecules to carry it closer. (n = 499, 42.6%) IV. The extracellular molecule collides randomly with other molecules. (n = 321, 27.4%)	1170	321 (27.4)	849 (72.6)
F	An extracellular molecule can change direction on its own. (True or False)	849	354 (62.9)	315 (37.1)
G	If extracellular molecules move via random collisions, what determines the chance of a binding event occurring between one of these molecules and a complementary receptor? I. If the binding event is required for cell function, it will happen regardless of other factors. II. The chance of the binding event occurring is determined by other factors, such as concentration and temperature.	321	272 (84.7)	49 (15.3)
H	In general, a large molecule (e.g., protein) has a more direct path of motion, whereas a small molecule (e.g., carbon dioxide or water) has a more random path of motion. (True or False)	849	348 (41.0)	501 (59.0)
I	A molecule’s path of motion is more direct when it has been activated (e.g., by phosphorylation), whereas its path of motion is more random when it is inactive. (True or False)	849	210 (24.7)	639 (75.3)
J	Inside a cell, macromolecules (e.g., proteins) are densely crowded, so much so that the average distance between two macromolecules is typically less than the width of a single macromolecule. (True or False)	1170	610 (52.1)	560 (47.9)
K	Inside a cell, water and other molecules are all in contact with each other; therefore, empty space does not dictate the direction of diffusion. (True or False)	1170	582 (49.7)	588 (50.3)
L	In the case of simple diffusion across a permeable membrane, once solute molecules reach an equilibrium, they cease to cross the membrane. (True or False)	1170	510 (43.6)	660 (56.4)
M	A drop of dye is placed in some water. The water, acting as a solvent, diffuses into the dye in the same way as the dye, acting as a solute, diffuses into the water. (True or False)	1170	667 (57.0)	503 (43.0)

confidence in both correct answers (Supplemental Tables 25 and 27, last row) and incorrect answers/misconceptions (Supplemental Tables 26–28, last row) differed significantly between all educational levels, with first-years expressing the lowest confidence and third-years expressing the highest confidence. For instance, first-year students were, on average, reasonably confident (51.39%) in their correct answers, while third-year students were very confident (72.94%) in their correct answers ($\eta^2 = 0.040$). Similarly, first-years were reasonably confident (50.30%) and third-years were somewhat more

confident (68.56%) in their misconceptions ($\eta^2 = 0.025$). Average confidence for misconceived answers was not significantly different than average confidence for correct answers in any student group (Supplemental Table 29).

At the level of individual MCAA items, confidence on correct responses was significantly higher among third-years than first-years for every item, while third-years also had higher confidence than second-years on the majority of items (Supplemental Tables 25 and 27). For those who answered with a misconception, the results show participants’ confidence is

TABLE 5. Pretest and posttest total misconceptions (as measured by the MCAA, maximum misconceptions = 11) in E2, stratified by biology course level

	n	Pretest			Posttest		
		Minimum, maximum	Mean	SD	Minimum, maximum	Mean	SD
First year	689	0, 11	5.96	2.24	0, 10	5.60	2.33
Second year	158	1, 10	5.97	2.15	0, 10	5.68	2.11
Third year	34	0, 10	5.00	2.88	0, 9	4.74	2.51

TABLE 6. Total misconceptions (as measured by the MCAA, maximum misconceptions = 11) at the beginning of the semester in 2015, 2016, and 2017 in small repeated-measures sample in E2, stratified by biology course level

	n	2015			2016			2017		
		Minimum, maximum	Mean	SD	Minimum, maximum	Mean	SD	Minimum, maximum	Mean	SD
First year (in 2015)	21	2, 9	5.57	2.23	0, 8	4.86	2.39	1, 9	4.33	2.69
Second year (in 2015)	7	2, 10	6.17	2.43	3, 10	5.86	2.27	2, 10	6.29	3.40

stronger at higher educational levels in statements A, E, H, and I (Supplemental Tables 26 and 28). For example, for statement A, third-years were 76.19% certain that a molecule “tries to move toward a complementary receptor,” while first-years were only 53.37% certain ($\eta^2 = 0.03$). The largest effects are seen in statement E (regarding the mechanism of motion), for which novice first-year students were 43.90% confident in their misconception, and second-years were 51.25% confident, whereas the more advanced third-year group was, on average, 76.09% confident (first to third: $\eta^2 = 0.06$; second to third: $\eta^2 = 0.08$).

Over the course of a semester. For second-year students, average confidence in correct responses increased significantly over the course of the semester by ~7.5% ($r_{cc} = 0.40$), while that of first-years (though significant) increased by only 1.6% ($r_{cc} = 0.09$), and third-years by 2.9% (not significant) (Supplemental Table 30). Similarly, second-year students’ confidence in their misconceptions increased significantly by 8.1% ($r_{cc} = 0.41$), while nonsignificant increases were seen for first-year (1.2%) and third-year (3.4%) students (Supplemental Table 31).

At the level of individual MCAA statements—and selecting only those students who retained the misconception from the beginning to the end of the semester—we observe a small but significant increase in confidence levels for misconception A (molecules try to find binding partners) in first-year ($r_{cc} = 0.16$) and second-year ($r_{cc} = 0.37$) students, as well as in misconception E (mechanism of molecular motion) in first-year ($r_{cc} = 0.18$) and second-year ($r_{cc} = 0.43$) students. Finally, we see a decrease in confidence in misconception L (diffusion/equilibrium) in first-year students ($r_{cc} = 0.23$) over the course of the semester. Participants’ confidence in all other misconceptions, across years, did not change significantly over the course of the semester (Supplemental Table 32).

Q4. How Did the Rewording/Restructuring of the MCAA Influence Students’ Responses?

The chi-square results for the comparisons between E1 and E2 are displayed in Supplemental Table 33; paired final and pilot MCAA statements can be viewed side by side in Supplemental Table 9. The rewording and/or reorganization of five statements (E2: C, E, J, K, and M) proved to have small to moderate effects on whether students responded with a misconception. Misconception rates were significantly reduced in C, J, K, and M. For statement C, this change can be linked to the sequential reordering of the statement; all participants in E2 were exposed to C, while only those who responded with a misconception to statement A were exposed to C (formerly D) in E1. In statements J, K, and M, we attribute the reduction in misconceptions largely to the clearer wording of the questions. Statement J, containing the phrase “empty space” in reference to the cytoplasm, was

reworded to explicitly address the crowded nature of cellular environments at the *macromolecular* level. This rewording resulted in fewer reported misconceptions overall. Similarly, in statement K, the “empty space” problem is addressed by explicitly identifying the presence of water and other small molecules that are in contact with one another; therefore, “empty space” could be interpreted only as a void (lack of molecules), even if the student would originally have referred to the cytoplasm as “empty.” Finally, statement M clearly defines the activity of solvent and solute molecules in diffusion, so that the student can make a more accurate decision as to whether the statement fits his or her current perception.

Contrastingly, we see increased misconceptions in statement E (mechanism of motion; formerly K in E1). This statement was identified as a core misconception in our focus group and 1) was moved up the structure of the survey and 2) had an additional answer option added in response to the focus group findings (i.e., molecules are carried by “helpers”); therefore, a greater proportion of students were exposed to the question. In comparing the distribution of participants’ chosen responses to this statement in E1 (statement J in Supplemental Table 7) and in E2 (statement E in Supplemental Table 22), we observed that, in E1, students evenly selected that 1) molecules are self-propelled and 2) molecules are released from a source with a correct trajectory; contrastingly, in E2, the large majority of students who responded with a misconception (58.8%) selected the newly provided option: “the extracellular molecule uses other ‘helper’ molecules to carry it closer [to a binding partner].”

Similarly, statement D (formerly E) included a new option derived from focus group findings. Statement D attempts to characterize the misconception identified in statement C, that a molecule knows a physical location of a receptor. While the most popular rationale selected to justify this thinking in E1 was “through interactions when it is close to the receptor,” the new option introduced in E2, “it receives a message from elsewhere (e.g., from nucleus),” drew the highest number of responses.

E2 Discussion

Main Findings. This study implemented the Molecular Concepts Adaptive Assessment (MCAA) across introductory through advanced biology courses and in three consecutive academic years. The results show that undergraduate biology students hold robust misconceptions regarding the emergent nature of molecular 1) motion, 2) interactions, and 3) environments, and that these misconceptions do not resolve themselves over time; this is an issue that has been well documented in the existing literature (Chi and Roscoe, 2002; Slotta and Chi, 2006; Chi et al., 2012; Garvin-Doxas and Klymkowsky, 2008; Momsen et al., 2010; Jenkinson and McGill, 2012, 2013).

The current study extends the literature by further characterizing these misconceptions through adaptive assessment and by determining how firmly embedded these beliefs are with an analysis of confidence levels. For example, only about one out of every four students could identify random collisions as the primary mechanism of molecular motion toward a binding partner (statement E), with the majority of others preferring to answer that molecules must be carried to target destinations by “helper” molecules. This finding reflects our focus group discussions from E1, wherein students acknowledged that molecules move randomly but this line of reasoning broke down when students considered random motion in the context of cellular processes. While this misconception is held in equal proportions by first-, second-, and third-year students, their confidence in their incorrect answers increases over the course of the semester and in each increasingly advanced biology course. We were further able to characterize the nature of this misconception through statements F, H, and I; while students are less likely to believe that molecules can independently change direction, they believe prevalently that the activation or phosphorylation of molecules corresponded with a state change from moving randomly to moving more directly, a belief that, again, increases in certainty in students enrolled in more advanced courses. A slight improvement we did see over the term, in relation to autonomous movement of molecules (F) and diffusion of molecules (L), is likely explained by the subject matter being addressed in second- and third-year curricula (diffusion, in particular, is covered in detail in the second year). Nevertheless, a surprising number of upper-level students struggle with understanding this material. As well, the finding of a weak negative correlation with science literacy and bio-literacy scores suggests that those students who hold the greatest number of misconceptions experience difficulty across the board.

That confidence in misconceptions increases with exposure to undergraduate education, while the rates of misconceptions themselves do not decrease significantly, suggests that the way in which we teach cellular/molecular biology—and/or the educational media that is used (Jenkinson and McGill, 2012; Cooper *et al.*, 2013)—reinforces students’ misunderstandings of molecular agency. Alternatively, these findings could suggest that students become more confident in their intuitive reasoning, because, as they progress through their undergraduate careers, their courses do not address these misconceptions head-on. This reasoning is supported by the finding that confidence in correct answers on the MCAA significantly increased in more advanced courses, as well as over a single semester, and is highly correlated to confidence in incorrect answers (i.e., students who are confident in their correct answers are also confident in their misconceptions).

A rather striking misconception, revealed through E1’s focus group discussion and confirmed in the survey responses in E2, was undergraduates’ prevalent belief that the nucleus is able to communicate with molecules or direct them in some way to binding partners. Approximately 56% of students who agreed that molecules know the location of binding partners indicated that molecules receive this knowledge through messages from the nucleus. Students may be applying their understanding of human anatomy, where the brain is able to send electrochemical signals in a seemingly direct manner to control parts of the body. In other words, they apply the schema of a well-known

direct or sequential system (at the macro level) to an inherently emergent system (at the micro level; Chi, 2005; Chi and Ohlsson, 2005; Chi *et al.*, 2012). This may be an important finding for how we discuss the functions of the nucleus in the classroom, to break the association between micro-level emergent systems and macro-level sequential referents.

Students’ belief in molecular agency is further exemplified in their answers relating to simple diffusion concepts. Our results support previous research by Meir *et al.* (2005), who found that a large proportion of college students held misconceptions regarding the equivalency of solute/solvent roles, directed motion of solute molecules toward lower concentrations, static equilibrium, and other concepts in osmosis and diffusion. Diffusion is a subject that is covered in high school and is revisited in first-year university; while it may be understandable for novice first-year students to believe that molecules cease to move across semipermeable membranes once equilibrium has been reached (misconception L), it is surprising that half of all third-year participants continue to hold this misconception and are “fairly confident” in their understanding of this process. Even so, we observed a significant 6.5% drop in misconception L over the course of the semester, and we also observed a decrease in confidence among first-year students who retained the misconception, suggesting that they no longer hold the misconception as firmly as before. Nevertheless, given that a similar proportion of students hold this misconception across 3 years of study, this illustrates a need for the subject matter to be revisited and refreshed in the curricula of more advanced courses.

The implications of these findings are potentially far-reaching. First, they can serve as a guide for postsecondary biology teachers when addressing this complex subject matter in their lectures; the MCAA itself could be used by teachers as a formative assessment throughout the term to ensure that conceptual change is affected (Kaufman *et al.*, 2002). Second, characterization of these misconceptions is critical for designing learning media that does not perpetuate misunderstandings. Dynamic representations of molecular interactions (animations and sequential illustrations) typically portray cellular processes in a linear and directed manner. Partly, in the case of static illustrations, this is a limitation of the medium. However, by understanding students’ preconceived notions of molecular agency, developers of educational media can be more careful in the design of representations, so that misconceptions are not reinforced by oversimplified visualizations (Mammino, 2008; Jenkinson and McGill, 2012; Cooper *et al.*, 2013). Finally, a detailed characterization of misconceptions is much needed in order to develop interventions that help students resolve their understanding. This study has shown, in several different ways (i.e., 1, comparison between introductory, intermediate, and advanced biology courses; 2, change from pre- to postsemester; and 3, longitudinal change over 3 years with repeated-measures participants), that misconceptions regarding molecular emergence, for the most part, persist over time and may even be reinforced through standard educational practices; innovative means of targeting these misconceptions are needed. Interactive molecular simulations and digital games have successfully been used to foster conceptual change in undergraduate biology students by visually modeling an emergent system and supporting students in confronting their misconceptions (Meir *et al.*, 2005;

Slotta and Chi, 2006; Persson *et al.*, 2007; Gauthier and Jenkinson, 2017). While no one type of visualization might effectively convey all critical aspects of emergent molecular systems, multiple representations might be used in tandem to support students' development of more accurate mental models (Ainsworth and VanLabeke, 2004; Ainsworth, 2006).

Limitations and Future Directions. Our study benefited from a large sample of students and longitudinally followed a small proportion of them over the course of their undergraduate careers, which, to the best of our knowledge, has not previously been done. Our analyses of the MCAA responses show the validity of the assessment: significant differences in results did not exist between 2015 and 2016 distributions. They also highlighted the accuracy of the assessment: responses in E2 reflect focus group discussions in E1; and the proportional change of misconception responses between E2 and E1 is consistent with the changes made to the structure and phrasing of the survey in light of the focus group discussions.

Despite these strengths, the present research has some limitations that may reduce the generalizability of our findings. First, the survey was only distributed among students at a single university in Canada and is not necessarily reflective of the undergraduate life sciences experience as a whole. Additionally, student enrollment in biology decreases at advanced levels of study, and our advanced-learner group (third-years) was thus much smaller than in our first- and second-year groups, limiting the generalizability of our third-year participant findings. Furthermore, we suffered significant sample attrition in our longitudinal cohort, which introduces bias; while our analyses suggest that individuals lost to attrition had similar misconceptions to the retained group, we cannot rule out other factors (e.g., course engagement, academic performance) that may have played a role in continued participation.

Future research should focus on collecting more data from advanced-learner groups (third- and fourth-year undergraduates) and across multiple university sites, so that the results can be broadly generalized. Also, distributing the survey among graduate students at various stages of study might give us a complete picture about how and when (or if) these misconceptions are eventually resolved. In addition, the model of adaptive assessment that was used to characterize student understanding could be extended to examine student understanding related to many other so-called threshold concepts in life sciences education.

CONCLUSION

Misconceptions regarding the emergent nature of molecular environments are deeply ingrained and resistant to change in undergraduate biology education. In this study, we developed and distributed an adaptive assessment designed to provide a detailed characterization of these misconceptions. Our results confirmed that students hold persistent misconceptions regarding molecular agency, environments, and diffusion throughout first-, second-, and third-year university courses. We further identified that confidence or certainty in these misconceptions, particularly regarding molecular agency and directed molecular motion, increases over time. A primary finding was that students' understanding of random/Brownian motion is highly contextual; they have difficulty applying that understanding to cellular

processes. Such characterization is important when developing learning media and targeted interventions. Future studies should distribute the assessment to other universities, so that the results can be generalized across undergraduate- and graduate-level study, in order that we can better understand how, when, and under what circumstances conceptual change occurs.

ACKNOWLEDGMENTS

This research was supported in part by grants NSF #DUE-1220512 from the National Science Foundation and SSHRC #SIG-13/14 from the Social Sciences and Humanities Research Council. We thank the faculty (Professors Anderson, Arts, Chatfield, Espie, and Rawle) and students of the undergraduate biology program at the University of Toronto Mississauga for participating in this research over the past several years. This research was approved by the University of Toronto's Office of Research Ethics.

REFERENCES

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198. doi: 10.1016/j.learninstruc.2006.03.001
- Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, 14(3), 241–255. doi: 10.1016/j.learninstruc.2004.06.002
- Andrews, T., Leonard, M. J., Colgrove, C., & Kalinowski, S. T. (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE—Life Sciences Education*, 10(4), 394–405.
- Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, 14(2), 161–199. doi: 10.1207/s15327809jls1402_1
- Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In Vosniadou, S. (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale, NJ: Erlbaum.
- Chi, M. T. H. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In Vosniadou, S. (Ed.), *International handbook of research on conceptual change* (pp. 49–70). New York: Routledge.
- Chi, M. T. H., & Ohlsson, S. (2005). Complex Declarative Learning. In Holyoak, K. J., & Morrison, R. G. (Eds.), *Cambridge handbook of thinking and reasoning* (p. 59). New York: Cambridge University Press.
- Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In Limon, M. & L. Mason, M. (Eds.), *Reconsidering conceptual change: Issues in theory and practice*. (pp. 3–27). Dordrecht, Netherlands: Kluwer Academic.
- Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, 36(1), 1–61. doi: 10.1111/j.1551-6709.2011.01207.x
- Cooper, S., Khatib, F., & Baker, D. (2013). Increasing public involvement in structural biology. *Structure*, 21(9), 1482–1484. doi: 10.1016/j.str.2013.08.009
- Couch, B. A., Wood, W. B., & Knight, J. K. (2015). The molecular biology capstone assessment: A concept assessment for upper-division molecular biology students. *CBE—Life Sciences Education*, 14(1), ar10. doi: 10.1187/cbe.14-04-0071
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, 13(3), 478–492. doi: 10.1187/cbe.13-10-0204
- Ellis, R. J. (2001). Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, 26(10), 597–604. DOI: 10.1016/S0968-0004(01)01938-7
- Fisher, K. M., Williams, K. S., & Anderson, D. L. (2012). *Concept inventories/ conceptual assessments in biology (CABs): An annotated list*. Retrieved December 13, 2017, from www.sci.sdsu.edu/CRMSE/files/Concept_Inventories_in_Biology_20110325.pdf

- Garvin-Doxas, K., & Klymkowsky, M. W. (2008). Understanding randomness and its impact on student learning: Lessons learned from building the Biology Concept Inventory (BCI). *CBE—Life Sciences Education*, 7(2), 227–233. doi: 10.1187/cbe.07-08-0063
- Gauthier, A., & Jenkinson, J. (2015). Game design for transforming and assessing undergraduates' understanding of molecular emergence (pilot). In Munkvold, R., & Kolas, L. (Eds.), *Proceedings of the 9th European conference on games based learning* (pp. 656–663). Steinkjer, Norway: Academic Conferences and Publishing International.
- Gauthier, A., & Jenkinson, J. (2017). Serious game leverages productive negativity to facilitate conceptual change in undergraduate molecular biology: A mixed-methods randomized controlled trial. *International Journal of Game-Based Learning*, 7(2), 20–34. doi: 10.4018/IJGBL.2017040102
- Guzmán, E., & Conejo, R. (2004). A model for student knowledge diagnosis through adaptive testing. In Lester, J., Vicari, R., & Paraguacu, F. (Eds.), *Intelligent tutoring systems* (Vol. 6, pp. 12–21). Dordrecht, Netherlands: Springer.
- Guzmán, E., & Conejo, R. (2015). Measuring misconceptions through item response theory. In Conati, C., Heffernan, N., Mitrovic, A., & Verdejo, F. M. (Eds.), *Artificial intelligence in education: 17th international conference, AIED 2015, Madrid, Spain, June 22–26, 2015. Proceedings* (pp. 608–611). Dordrecht, Netherlands: Springer International. doi: 10.1007/978-3-319-19773-9_73
- Holm, S. (1979). A simple sequential rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- Huang, Y. M., Lin, Y. T., & Cheng, S. C. (2009). An adaptive testing system for supporting versatile educational assessment. *Computers and Education*, 52(1), 53–67. doi: 10.1016/j.compedu.2008.06.007
- IBM Corporation. (2017). IBM SPSS Statistics for Mac/Intosh, Version 25.0. Armonk, NY.
- Jenkinson, J., & McGill, G. (2012). Visualizing protein interactions and dynamics: Evolving a visual language for molecular animation. *CBE—Life Sciences Education*, 11, 1–8.
- Jenkinson, J., & McGill, G. (2013). Using 3D animation in biology education: Examining the effects of visual complexity in the representation of dynamic molecular events. *Journal of Biocommunication*, 39(2), 42–49.
- Kalas, P., O'Neill, A., Pollock, C., & Birol, G. (2013). Development of a meiosis concept inventory. *CBE—Life Sciences Education*, 12(4), 655–664. doi: 10.1187/cbe.12-10-0174
- Kaufman, R., Watkins, R., & Guerra, I. (2002). Getting valid and useful educational results and payoffs: We are what we say, do, and deliver. *International Journal of Educational Reform*, 11(1), 77–92.
- Klymkowsky, M. W., & Garvin-Doxas, K. (2008). Recognizing student misconceptions through Ed's Tools and the Biology Concept Inventory. *PLoS Biology*, 6(1), e3. doi: 10.1371/journal.pbio.0060003
- Klymkowsky, M. W., Underwood, S. M., & Garvin-Doxas, R. K. (2010). *Biological Concepts Instrument (BCI): A diagnostic tool for revealing student thinking*. arXiv:1012.4501v1.
- Mammino, L. (2008). Teaching chemistry with and without external representations in professional environments with limited resources. In Gilbert, J. K., Reiner, M., & Nakhleh, M. (Eds.), *Visualization: Theory and practice in science education* (pp. 155–185). Dordrecht, Netherlands: Springer.
- Meir, E., Perry, J., Stal, D., Maruca, S., & Klopfer, E. (2005). How effective are simulated molecular-level experiments for teaching diffusion and osmosis? *Cell Biology Education*, 4(3), 235–248. doi: 10.1187/cbe.04-09-0049
- Miller, J. D. (1998). The measurement of civic scientific literacy. *Public Understanding of Science*, 7, 203–223.
- Momsen, J. L., Long, T. M., Wyse, S. A. S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE—Life Sciences Education*, 9(4), 435–440. doi: 10.1187/cbe.10
- Nehm, R., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57, 263–272.
- Newman, D. L., Snyder, C. W., Fisk, J. N., & Wright, L. K. (2016). Development of the Central Dogma Concept Inventory (CDCI) assessment tool. *CBE—Life Sciences Education*, 15(2), ar9. doi: 10.1187/cbe.15-06-0124
- Odom, A. L. (1995). Secondary & college biology students' osmosis misconceptions about diffusion and osmosis. *American Biology Teacher*, 57(7), 409–415.
- Persson, P. B., Cooper, M. D., Tibell, L. A. E., Ainsworth, S., Ynnerman, A., & Jonsson, B. H. (2007). Designing and evaluating a haptic system for biomolecular education. In *Proceedings of the IEEE virtual reality conference (VR 2007)* (pp. 171–178). New York: IEEE. Retrieved January 18, 2019, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.138.4869&rep=rep1&type=pdf>
- Robic, S. (2010). Ten common misconceptions about protein structure, folding, and stability. *CBE—Life Sciences Education*, 9, 189–195. doi: 10.1187/cbe.10
- Rundgren, C.-J., & Tibell, L. A. E. (2009). Critical features of visualizations of transport through the cell membrane—An empirical study of upper secondary and tertiary students' meaning-making of a still image and an animation. *International Journal of Science and Mathematics Education*, 8(2), 223–246. doi: 10.1007/s10763-009-9171-1
- Schönborn, K. J., & Bögeholz, S. (2009). Knowledge transfer in biology and translation across external representations: Experts' views and challenges for learning. *International Journal of Science and Mathematics Education*, 7(5), 931–955. doi: 10.1007/s10763-009-9153-3
- Shi, J., Wood, W. B. W. B., Martin, J. M. J. M., Guild, N. A. N. A., Vicens, Q., & Knight, J. K. J. K. (2010). A diagnostic assessment for introductory molecular and cell biology. *CBE—Life Sciences Education*, 9(4), 453–461. doi: 10.1187/cbe.10-04-0055
- Slotta, J. D., & Chi, M. T. H. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, 24(2), 261–289. doi: 10.1207/s1532690xci2402_3
- Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The Genetics Concept Assessment: A new concept inventory for gauging student understanding of genetics. *CBE—Life Sciences Education*, 7(1), 422–430. doi: 10.1187/cbe.08
- Tibell, L. A. E., & Rundgren, C.-J. (2010). Educational challenges of molecular life science: Characteristics and implications for education and research. *CBE—Life Sciences Education*, 9(1), 25–33. doi: 10.1187/cbe.08-09-0055
- van der Linden, W., & Glas, C. (2000). *Computerized adaptive testing: Theory and practice*. Dordrecht, Netherlands: Kluwer Academic.