

Development and Validation of the Microbiology for Health Sciences Concept Inventory[†]

Heather M. Seitz^{1*}, Rachel E. A. Horak², Megan W. Howard³, Lucy W. Kluckhohn Jones⁴, Theodore Muth⁵, Christopher Parker⁶, Andrea Pratt Rediske⁷, and Maureen M. Whitehurst⁸

¹Johnson County Community College, Overland Park, KS 66220; ²American Society for Microbiology, Washington, DC 20036; ³Battelle Memorial Institute, Columbus, OH 43201; ⁴Santa Monica College, Santa Monica, CA 90405; ⁵CUNY Brooklyn College, Brooklyn, NY 11210; ⁶Texas Wesleyan University, Fort Worth, TX 76105; ⁷University of Central Florida, Orlando, FL 32816; ⁸Trident Technical College, Charleston, SC 29406

Identifying misconceptions in student learning is a valuable practice for evaluating student learning gains and directing educational interventions. By accurately identifying students' knowledge and misconceptions about microbiology concepts, instructors can design effective classroom practices centered on student understanding. Following the development of ASM's Curriculum Guidelines in 2012, we developed a concept inventory, the Microbiology for Health Sciences Concept Inventory (MHSCI), that measures learning gains and identifies student misconceptions in health sciences microbiology classrooms. The 23-question MHSCI was delivered to a wide variety of students at multiple institution types. Psychometric analysis identified that the MHSCI instrument is both discriminatory and reliable in measuring student learning gains. The MHSCI results correlated with course outcomes, showing the value of using the instrument alongside course level assessments to measure student learning. The MHSCI is a reliable and efficient way to measure student learning in microbiology and can be used both as a faculty development tool and an effective student assessment tool.

INTRODUCTION

Beginning with the Force Concept Inventory in 1992 (1), the value of identifying student misconceptions has been well studied. Student misconceptions can be an incredibly difficult barrier to overcome (2), often requiring interventions to uncover (3, 4). The first step in correcting student misconceptions is to expose them to both the faculty member and the student. Concept inventories are research-driven, validated instruments designed to measure expert-like thinking on specific topics (5). Initially, concept inventories were developed for the physical sciences; however, since the initial inventory, multiple concept inventories have been developed for biology concepts, including general biology (6), biology experimental design (7), genetics (8),

host-pathogen interactions (9), meiosis (10), molecular biology (11), the lactose operon (12), and biomechanics (13). While there are now multiple biology concept inventories available, there are currently no concept inventories that align with the curriculum for health sciences microbiology taught to a variety of students (pre-nursing, pre-dental hygiene, pre-pharmacy, veterinary technology).

In 2012, the ASM Curriculum Guidelines for Undergraduate Microbiology were released (14), which include core concepts and fundamental statements for microbiology and provide a framework for content and skills in an undergraduate microbiology classroom. Although faculty generally found the ASM Curriculum Guidelines useful, the microbiology community called for assessment resources to assist in the implementation of these curriculum guidelines in their classrooms (15). This task force, composed of health science faculty, has developed a valuable assessment tool to measure whether students were meeting these curriculum guidelines specifically focused on the health science microbiology curriculum.

Early in the process of creating an assessment instrument for the ASM Curriculum Guidelines, a diverse group of microbiology faculty was tasked with determining learning outcomes important for the undergraduate

*Corresponding author. Mailing address: Science Division, Johnson County Community College, 12345 College Blvd., Overland Park, KS 66210-1299. Phone: 913-469-8500 (4375).

E-mail: hseitz@jccc.edu.

Received: 18 March 2017, Accepted: 28 July 2017, Published: 30 October 2017.

[†]Supplemental materials available at <http://asmscience.org/jmbe>

classroom. It became apparent early in the process that the level of prerequisite knowledge and the focus on clinical applications was decidedly different between a microbiology course taken as part of a biology major at a four-year institution and one that was taken in preparation for an allied health field. As a result, the initial group of faculty was divided into two task forces, and two separate but aligned concept inventories were created. Following the steps outlined by Adams and Wieman (16), we created a concept inventory directly aligned with the ASM Curriculum Guidelines, with the guidance of health sciences faculty from professional programs.

Microbiology and biology faculty developed two separate but aligned concept inventories: the Microbiology Concept Inventory (MCI) is intended for a microbiology course taken with rigorous prerequisites and as part of a biology major course track. The MCI is described and published in this same issue. The MHSCI described in this article is intended for nonmajors and allied health students.

METHODS

The process of creating the concept inventory was guided by Adam and Wieman's work (16). Steps followed included 1) determining the microbiology concepts most important to faculty experts, 2) identifying student thinking on concepts, 3) creating an open-ended survey to probe student thinking more deeply, 4) developing a forced answer survey that measures student thinking, 5) carrying out validation interviews, 6) delivering the instrument, and 7) assessing results using statistical analyses.

Task force recruitment

Members of the task force were recruited through the ASM BiologyScholars listserv. Volunteers from a wide range of institution types and sizes were solicited. Task force members met in a series of web conferences and distance conversations. Participating institutions and task force members can be found in the supplemental materials. One task force meeting was held in person in March 2016 to assist in finalizing and validating the concept inventory.

Institutional Review Board (IRB) approval was obtained for all work involving students with the concept inventory questions. Task force members sought IRB approval at their own institutions and worked within the framework of their IRB guidelines to protect student privacy. This study was completed in compliance with human subject IRB 141007 (Johnson County Community College), IRB L.Mehlig@RockValleyCollege.edu (Rock Valley College), IRB 2014-1466, 2015-1272 (University of Wisconsin – Madison), IRB 20150601_044SCMC (Santa Monica College), IRB 2016-2750 (University of California – Irvine), and IRB 01-2015 (Trident Technical College). No human subject information was handled at Battelle.

Survey of professional program faculty

Initial work on the concept inventory began in summer 2014. To identify key concepts critical in health science professions, a survey was sent to health science professional faculty (Nursing, Dental Hygiene, Respiratory Therapy, and Pharmacy) to help determine which of the ASM Curriculum Guidelines fundamental statements were most important for their incoming students. The ASM Curriculum Guidelines (14) include 27 fundamental statements incorporating six core concepts that the microbiology community has agreed are essential and important for undergraduate students. For our particular population of students, those interested in healthcare fields, we felt that additional learning outcomes should be included. Therefore, we included fundamental statements from the Host-Pathogen Inventory (9) that deal directly with the immune system and host-pathogen interactions that were not covered in the ASM Curriculum guidelines. We also included additional immunology-related learning outcomes, as immunology is a common component of an allied health microbiology course and is not fully covered by the ASM Curriculum Guidelines. The survey asked faculty members to rate the importance of each fundamental statement to their curriculum using a Likert scale of 1 to 5, with 1 being “not important at all” and 5 being “very important.” Four different programs at three institutions were asked to complete the survey, and 35 completed responses were collected. Survey respondents fell into four self-identified program types (Fig. 1A). Fundamental statements deemed most important from survey respondents are shown in Figure 1B. Responses clearly showed that concepts related to environmental microbiology, metabolic pathways, and evolution were not deemed as important to the allied health professional faculty as key concepts related to host-pathogen interactions, immunology, cellular structure and function, and information flow and genetics (Fig. 1B). These data were used as a guide as the task force determined which concepts to include.

Development of learning outcomes

Upon reviewing the results of the faculty survey, task force members carefully considered the fundamental statements that were most often selected as highly important and narrowed the list of statements. Task force members also agreed not to include statements that were not selected as important by a majority of the faculty respondents. After narrowing down the list of fundamental statements the task force wrote 23 learning outcomes (Table 1).

Inventory development

The task force developed a series of T/F questions to address each learning outcome. Faculty on the task force wrote the T/F questions to align with the learning outcomes, and the questions were then reviewed and revised by the

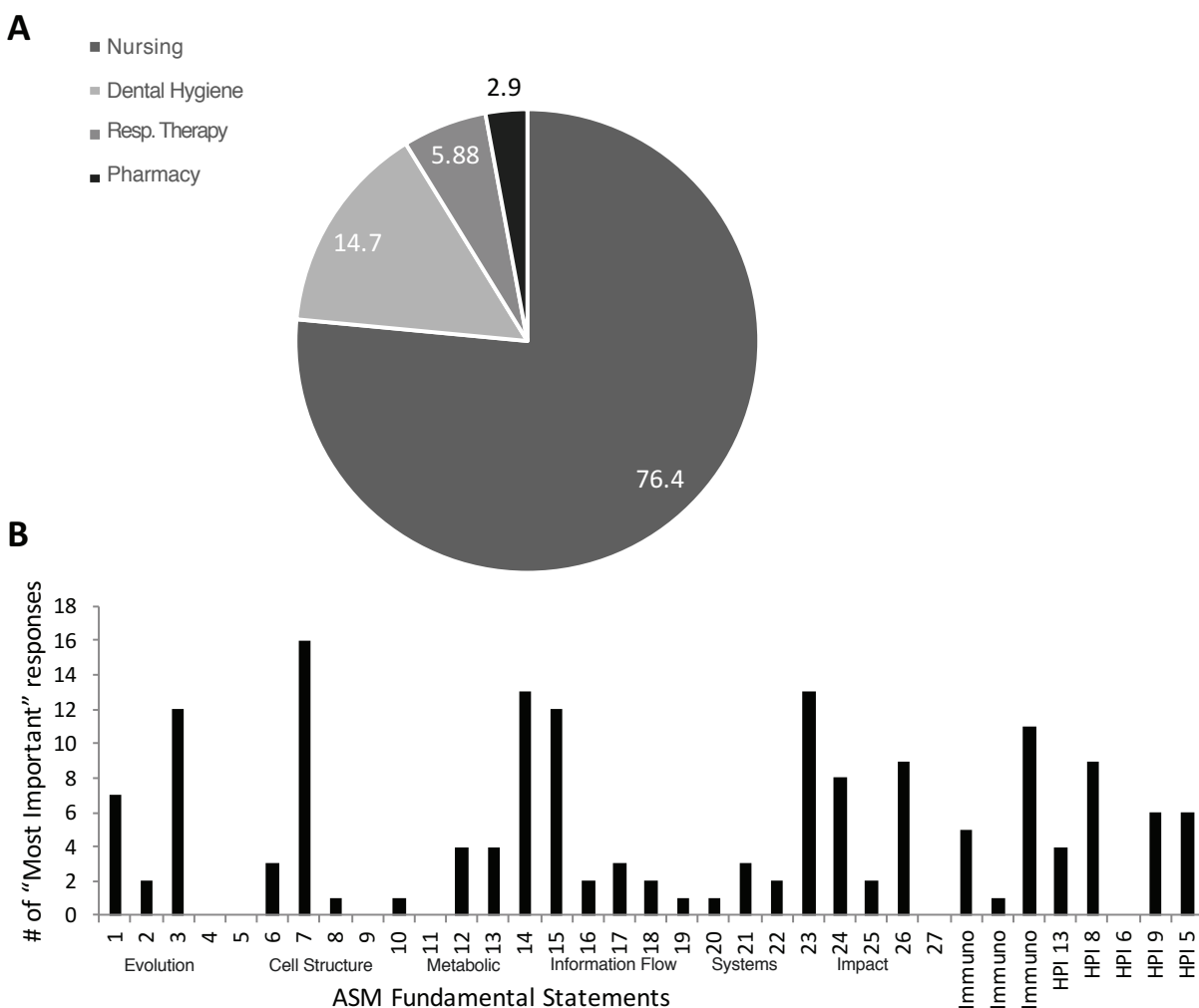


FIGURE I. Survey results from professional program faculty. A) Survey was sent to professional program faculty, and 35 faculty responded. The percentage of each faculty program is shown in the pie chart. B) Survey results from professional faculty on what the most important concepts were for students to understand for their professional program. Core concepts are written below the fundamental statement number for reference. Immuno = additional immunology related concepts; HPI = learning outcomes sourced from the Host Pathogen Interactions Concept Inventory (9); ASM = American Society for Microbiology.

entire task force. The completed 26 T/F questions were then delivered to students (T/F questions can be found in the supplemental materials). Students were surveyed using this list of questions, and their short responses explaining why they felt the statement was either true or false were collected. Initially, 119 responses were collected at three different institutions (sample student responses can be found in supplemental materials). These data were compiled, and three teams of three task force members each were assigned to review responses for each question and qualitatively code responses to identify emerging themes (17). To triangulate the qualitative coding, task force faculty discussed and developed agreed-upon consensus codes for each question. Using these consensus codes, each of the responses and student explanations to each question were recoded by three task force members to identify common student misconceptions, which were then mapped to specific

ASM fundamental statements. After consensus codes were applied, task force faculty discussed any items with divergent codes between the faculty members, and the most common incorrect student responses were determined. The three teams all had interrater reliability above 70% (Team 1 – 81%, Team 2 – 71.3%, and Team 3 – 88.9%) (18). The misconceptions identified are similar to those found by the task force developing the general concept inventory (19).

After reviewing and coding the responses to T/F questions, faculty tasked with coding each question generated a multiple-choice question (MCQ) addressing the learning outcome for that concept utilizing the most common incorrect student responses as distractors. Jargon was minimized, and student wording was incorporated into answer choices. The whole task force reviewed questions for accuracy and clarity. The initial multiple-choice concept inventory was delivered to 209 students at four

TABLE I.
MHSCI questions aligned with ASM and HPI fundamental statements and learning objectives.

MHSCI Question	Fundamental Statements	Learning Outcome
1	ASM 2	Give an example of a bacterial pathogen that evolved naturally or artificially to become attenuated (e.g., vaccine strains, intracellular pathogens, etc.).
2	ASM 3	Explain how public health policies (e.g., quarantine and vaccination) can alter epidemic/pandemic progression.
3	ASM 2, 3	Describe how mutations and horizontal gene transfer, together with selective pressure, can lead to a rise of antibiotic resistance or the spread of virulence mechanisms.
4	ASM 6	Relate the sizes of human cells, eukaryotic pathogen cells, bacterial cells and viruses using appropriate units of measurement for cell sizes.
5	ASM 7	Design a target for a new drug based on the structure of bacterial cells.
6	ASM 7, 8	Given the role of some specialized cell structures (such as capsules, flagella, fimbriae, spores, secretion systems) in pathogenesis, predict how gaining or losing the ability to make a functional version of the specialized structure might affect the organism's ability to be a pathogen.
7	ASM 10	Correlate the replication cycles of viruses to the types of infections caused and treatments administered to the host.
8	ASM 6	Determine the appropriate type of microscopy and staining to visualize a specific microorganism or structure (flagella stain, capsule stain, acid fast stain).
9	ASM 12	Describe how bacteria shift their metabolism in response to their environment (e.g., inside/outside of a host, oxygen concentration, presence/absence of other bacteria, presence/absence of substrates).
10	ASM 14	Given a particular situation, choose the best method (e.g., physical, chemical, biological, etc.) for controlling bacterial growth.
11	ASM 21	Give examples of how cells regulate their growth (metabolic activity) in response to other microbial cells.
12	ASM 15	Through an understanding of Griffith's classic experiment with rough and smooth cells, describe the relationship between capsule genes and virulence.
13	ASM 15	For a given point mutation, genetic insertion, or genetic deletion, describe a situation that would result in a non-functioning protein and one that would not.
14	ASM 14	Diagram/graph how altering environmental conditions (e.g., temperature, oxygen, nutrient levels, antibiotics, immune responses and disinfectant levels) will affect the growth of a population.
15	ASM 16	Explain how the organization of genes in an operon affects transcription in bacteria, compared with a single gene.
16	ASM 17	Give examples of how an external chemical signal can control gene expression.
17	ASM 18	Compare and contrast the host and viral enzymes needed by RNA, DNA, and retroviruses.
18	ASM 22	Describe how microorganisms interact with human hosts, including both positive and negative examples of these interactions.
19	ASM 26	Explain the importance of microbial fermentation products to food/beverage production (e.g., bread, cheese, yogurt, wine, beer, etc.).
20	ASM 24	List at least three reasons why microbes are important to life on earth (include probiotics, microbiota, and animal health).
21	HPI 12	Different types of innate and adaptive immune responses are required to combat extracellular and intracellular microorganisms.
22	Task Force Developed*	Describe how a vaccine can cause protection to a disease, yet not cause the disease (compare and contrast the immunological response to an attenuated versus inactivated vaccine).
23	HPI 11	Explain how the immune system of the body recognizes a pathogen early in an infection.

*Task Force Statement developed: The immune response to an attenuated/nonpathogenic form of a microorganism can still induce immunity.
ASM = ASM Curriculum Guidelines fundamental statement number; HPI = Host Pathogen Interactions Concept Inventory Concept Number.

institutions in 2015. In-person interviews with 15 students were also conducted using the “think-aloud” protocol (20). Interviews were used to understand which questions contained jargon that prevented students from answering the question, as well as to look for answer choices that were obvious due to grammar or word usage. Interview results were discussed among the entire task force, and adjustments were made as needed to questions and answer choices.

Data collected for each question of the concept inventory included answer choice and student reasoning. The distribution of answer choices was examined as was the difficulty and discrimination of each question (see below for calculations used). The goal was to see a wide distribution of answer choices for the pretest delivery of questions and a shift to the correct answer choice for the posttest delivery of the questions. Upon examining the original question data, the task force edited seven of the questions. Four questions were edited because they had very low difficulty as a pretest item (greater than 50% of students choosing the correct answer) and three questions were edited based on low discrimination values (less than 0.05). The updated assessment was delivered in fall 2015 as a posttest to 91 students at three different institutions. Further examination of question data revealed issues with discrimination that warranted further editing of three questions. Final review of the questions took place in an in-person meeting of the task force. A review panel was formed that included two faculty members who had not been involved in question development. The reviewers were tasked with looking at question accuracy, relevance, and alignment with learning outcomes and were asked to select the fundamental statements that each question aligned with. A complete discussion of the review panel’s responses was done in person and changes to four questions were made to ensure clarity and accuracy.

Inventory delivery

The final version of the concept inventory was delivered to 620 students as a pretest in summer/fall 2016 and as a posttest to 605 students in spring/summer/fall 2016. Faculty from six different institutions delivered the assessment as a pre- and posttest (two four-year institutions and four community colleges).

Psychometric analysis

Statistical analysis of the data was done using the guidelines and parameters found in Ding et al. (21). Briefly, for each question, the results were analyzed for item difficulty, discrimination, and reliability. Item difficulty is used to determine what percentage of students got each question correct. The lower the item difficulty score, the fewer students answered the question correctly, with an acceptable

threshold above 0.3. Item difficulty (P) was calculated using the formula:

$$P = \frac{N_1}{N},$$

where N_1 is the number of students who answered correctly and N is the total number of students. Item discrimination is used to determine how well a question separates the students who score well on the overall test from those students who do poorly on the overall test; acceptable threshold is above 0.3. A high discrimination score indicates that only students that did well on the overall test did well on that question. Item discrimination (D) was calculated using the formula:

$$\frac{N_H(\text{top } 25\%) - N_L(\text{bottom } 25\%) }{N/4},$$

where N_H is the number of students who answered the question correctly in the top 25% based on total score and N_L is the number of students who answered the question correctly in the bottom 25% based on total score. Similar to discrimination, the point-biserial coefficient (r_{pbs}) measures a single student’s answer in reference to their score on the whole test. However, r_{pbs} looks at the question as a predictor of score rather than the whole test score as a predictor for a question; the acceptable threshold for this calculation is above 0.2. The r_{pbs} was calculated to determine reliability of each test item using the formula:

$$r_{pbs} = \frac{\bar{X}_1 - \bar{X}}{\sigma_x} \sqrt{\frac{P}{1-P}},$$

where \bar{X}_1 is the average total score for those students who scored correctly on the item and \bar{X} is the average total score for the whole sample. σ_x is the standard deviation of the total score for the whole sample. Whole-test discrimination was calculated using the formula for Ferguson’s delta. Ferguson’s delta shows how broadly the total scores are distributed over all possible scores on the assessment. If the test is discriminatory, then there should be a broad distribution; the acceptable threshold is above 0.9. Ferguson’s delta was calculated using the formula:

$$\delta = \frac{N^2 - \sum f_i^2}{N^2 - N^2/(K+1)},$$

where f_i is the frequency of occurrence of each test score. Whole-test reliability was calculated using the Kuder-Richardson formula 21, which calculates whole test consistency; the acceptable threshold is above 0.7.

The Kuder-Richardson 21 formula was chosen because it is an appropriate measure of internal reliability for multiple-choice tests such as the MHSCI. The Kuder-Richardson formula 21 is primarily used for dichotomously scored items and generally yields the same results as Cronbach’s alpha for these types of assessments (22). Whole-test reliability was calculated using the formula:

$$r_{test} = \frac{K}{K-1} \left(1 - \frac{\sum P(1-P)}{\sigma_x^2} \right),$$

where K is the number of test items.

RESULTS

Inventory delivery

In order to ensure a representative student population, the MHSCI was intentionally delivered to a range of institution types (four two-year community colleges and two large research institutions) across a wide geographic range (United States: CA, KS, IL, SC, WI). Our sample was roughly equally distributed with respect to institution type, with 287 students surveyed from two-year institutions and 318 students from four-year institutions. The students who responded represented a very ethnically diverse population and also reflected the diversity seen in undergraduate institutions with regard to declared majors and programs (Fig. 2). The number of credit-hours completed before taking microbiology varied widely, with the majority of students completing four to five semesters of coursework (Fig. 2C). The ethnic diversity of our student population is comparable with national data statistics from 2014 (Fig. 2B) (23). There were also a large variety of declared majors and career programs, with the majority of students declaring nursing as their career program (Fig. 2A). While national averages for undergraduate biology graduation trends favor a 60% female versus 40% male distribution, our surveyed population showed a much larger disparity in the number of women versus men, with only 18% surveyed identifying as male and 82% identifying as female (24). In many cases, pre-professional microbiology courses are delivered at two-year community college institutions, and our surveyed student population reflects this fact.

Student learning gains

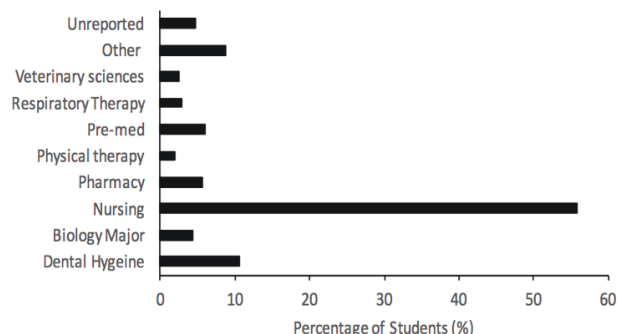
To test whether the concept inventory can measure student learning gains, we delivered the inventory as a pre- and posttest. On average, students scored 44.38% ($\pm 2.97\%$ SD) or 10 questions answered correctly out of 23 questions total on the pretest. There was significant improvement between the pretest and posttest scores ($55.73\% \pm 3.63\%$, or 13 of 23 questions answered correctly), with an average increase of 11.4% or 3 more correct questions answered (ANOVA single factor $F = 164.5$, $p < 0.005$) (Fig. 3A). Normalized learning gain for the MHSCI was 0.2, which puts the assessment in the low-gain category (25). The low-gain category is appropriate, as we are measuring misconceptions that have been shown repeatedly to be difficult to overcome in student thinking.

Item analysis

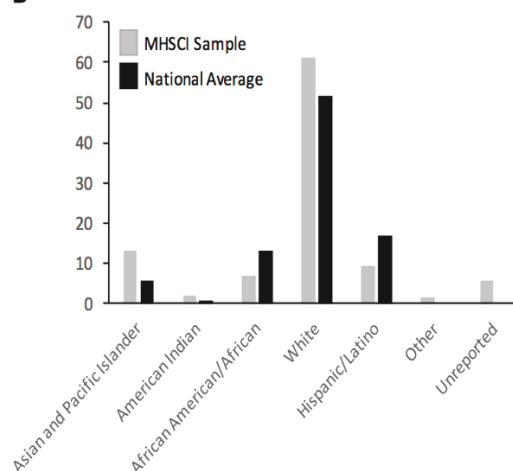
In order to properly evaluate each of the questions included in the concept inventory, we calculated the reliability, difficulty, and discrimination for each question (Figs. 3B, C, and D). The reliability score on average was $r_{pbs} = 0.34$, above the accepted value of 0.20 (24). Item difficulty

(P) was consistent across the questions, and averaging all of the questions gave a difficulty value of $P = 0.6$ which is within the range of accepted values (0.3 to 0.9). Difficulty values

A



B



C

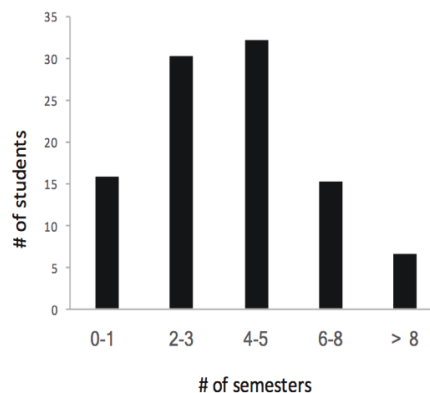


FIGURE 2. Demographic data of students surveyed. A total of 322 students completed the demographics questions within the MHSCI. Questions included age, sex, race, number of semesters, intended major, and familiarity with the English language. A) Declared majors or programs identified by students. B) The distribution of ethnicity, self-reported by the student sample. C) The number of semesters completed, self-reported by students in the sample population.

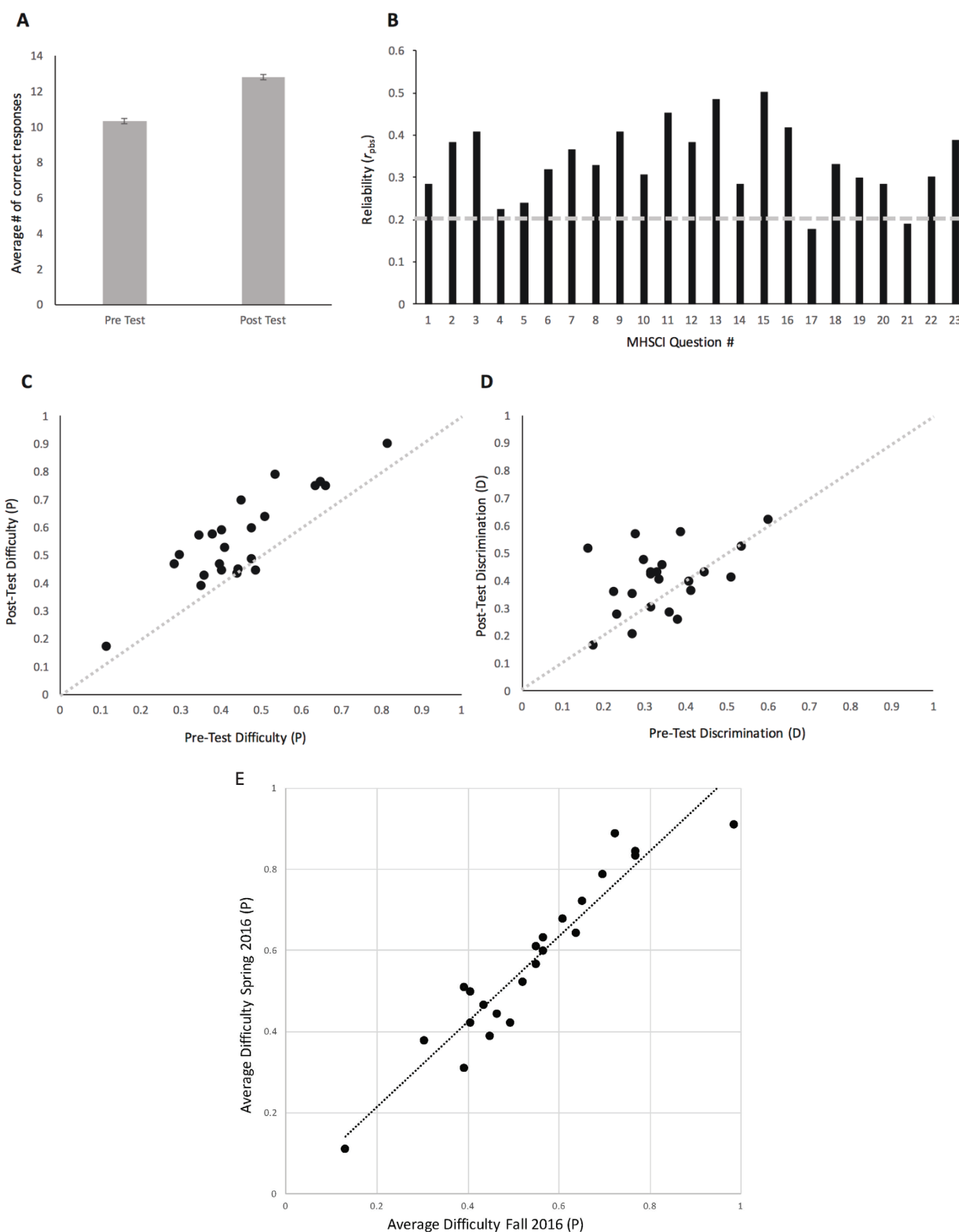


FIGURE 3. Psychometric analysis of MHSCI. A) Overall student score on the pretest and posttest. The error bars represent standard error. The increase in scores is significant ($p < 0.05$). B) Item reliability for each question was calculated using the point-biserial coefficient. The acceptable value of 0.20 is shown in the dashed gray line. C) Analysis of pretest and posttest difficulty. Each point represents a single question on the MHSCI. Points above the line show an increase in P, which demonstrates an increase in the number of students answering the question correctly on the posttest. D) Analysis of pretest and posttest discrimination; values above the dotted line demonstrate questions that showed increased discrimination between students scoring in the bottom 25% and the top 25%. E) Test-Retest measure – difficulty values from two semesters at the same institution are plotted by question. Each dot represents a question on the MHSCI.

on the pretest were lower than on the posttest for nearly every question (Fig. 3C). On average, the discrimination was 0.40, above the accepted value of 0.30 (Fig. 3D) (24). The above measures show that each item in the inventory is both valid and reliable.

Whole test reliability

In order to evaluate the reliability of the entire assessment, two measures were initially calculated: Ferguson's delta and the Kuder-Richardson formula 21. The MHSCI scored 0.96 on Ferguson's delta, which is well above the 0.90 threshold for reliability (26). The Kuder-Richardson accepted values are above 0.5; the calculated value for the assessment was KR_{21} 0.65. Test-retest measures are difficult to deliver in a single semester without an expectation of students remembering the questions. Due to this constraint and in order to test reliability, we utilized scores from two semesters of the same course at a single institution. Scores from two instructors were analyzed and compared in spring 2016 and fall 2016 (Fig. 3E). The scores from both semesters were comparable, with a Pearson's correlation of 0.95, indicating high correlation between the two scores. Looking at data from all institutions, we found that scores were not measurably different between institution type, gender, or age. Authors intend that the MHSCI will be valuable across institution types and student demographics, and this initial analysis suggests that this is true (Fig. 4).

In order to better understand how the results of the MHSCI correlate with student course success, one institution matched 55 individual posttest scores with course exam averages. Using Pearson's correlation coefficient, there was a significant positive correlation between exam average and posttest score $r_{50} = 0.63$ ($p < 0.02$) (Fig. 5).

DISCUSSION

The development of the MHSCI was concurrent with the development of a concept inventory for general microbiology students (27). Faculty who teach microbiology will find the two microbiology concept inventories developed useful and helpful in understanding their students' learning gains and misconceptions. The MHSCI described here is designed for use in courses with limited or no biology prerequisites and considered "nonmajors" biology. The MHSCI includes more questions written from a clinical perspective and fewer environmental and metabolism concepts than the more general Microbiology Concept Inventory (MCI). The MCI is designed for faculty teaching a majors course with significant biology prerequisite coursework. While there are distinct differences between the MCI and MHSCI, there are several ($n = 3$) shared questions (MHSCI 14 = MCI 12, MHSCI 15 = MCI 15, MHSCI 11 = MCI 21), which address microbial physiology, gene expression, and biofilms.

The MHSCI will be disseminated through the following website: <https://sites.google.com/view/mhsci/home>.

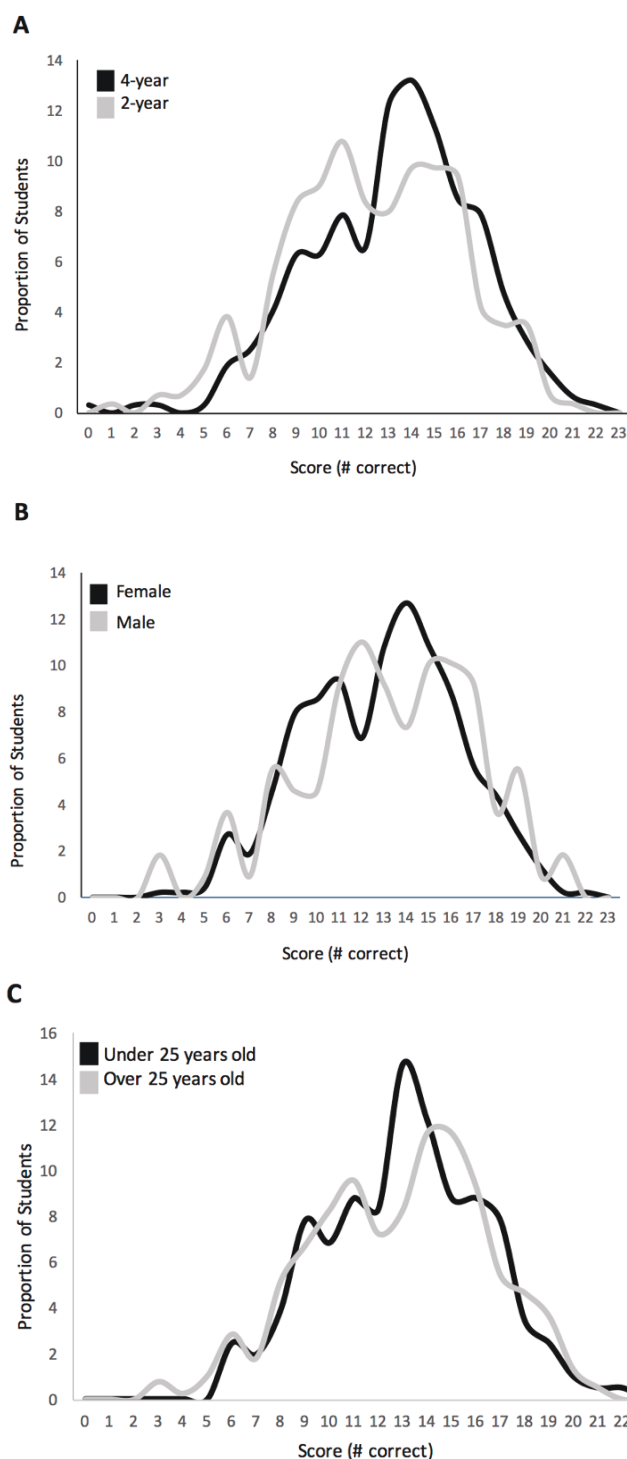


FIGURE 4. Correlation of MHSCI posttest score with institution type and demographic data. The curves represent the frequency of a given total score (0–23) for each group. A) Institution type, B) Gender, and C) Age.

Security of the assessment is important, and therefore the questions themselves can be requested through the corresponding author. The MHSCI is intended to help faculty measure learning gains over the course of a semester for the purpose of identifying specific student

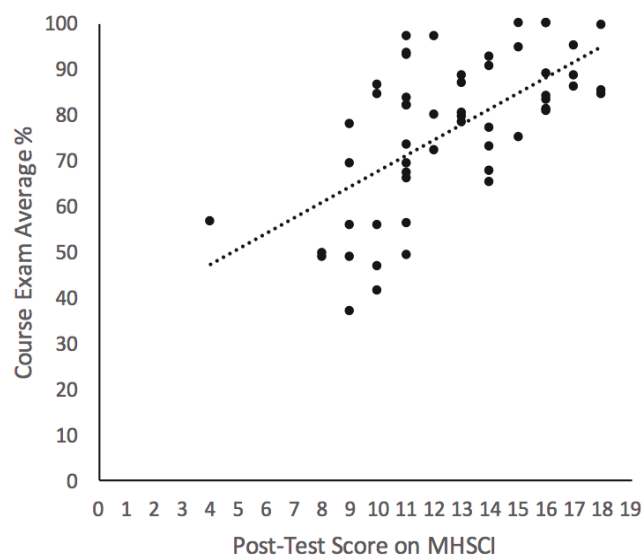


FIGURE 5. Correlation of MHSCI with course level assessment. Student posttest scores on the MHSCI were plotted along with their exam score average at the end of the course. Each dot represents a single student. The dashed line represents the best fit line for the data.

misconceptions and to guide curriculum development. Concept inventories are most valuable when used as a pretest and posttest with little discussion of the questions in between. Security and relevancy of the questions can be maintained by not providing students with an incentive that is based on their score on the MHSCI. For example, providing students with just five points of extra credit for completion of the pretest inventory, regardless of score, at the beginning of the course and five points of extra credit for completing the posttest inventory, in a course with 1,000 points total, has been effective, with over 90% of students completing both. Concept inventories measure expert-like thinking and are based on misconceptions; therefore it is not appropriate to use the score from the MHSCI as a course assessment grade.

Using the concept inventory as a faculty development tool can help guide the curriculum development process and help uncover topic-specific student misconceptions at the beginning of the course. One example is the faculty development of a microscope activity that utilizes images and size comparisons based on a low score for the MHSCI question concerning the relative size of bacteria. Further, faculty can identify misconceptions that remain at the end of the course, and discuss methods to help develop interventions to target course-specific misconceptions. The MHSCI is a rich assessment tool for health sciences faculty to understand their student misconceptions and will be a valuable asset to faculty learning communities.

The questions developed for this inventory are intended to be part of an iterative process. As we continue to collect data using the MHSCI with a larger student population, we anticipate continuing to revise questions in order to improve the inventory and ensure that concepts measured

are aligned with the most prevalent student misconceptions in the health sciences microbiology field.

SUPPLEMENTAL MATERIALS

- Appendix 1: Faculty and institution participation in the development of the MHSCI.
- Appendix 2: Flowchart of concept inventory development process.
- Appendix 3: The list of true/false statements used to elicit common student misconceptions.
- Appendix 4: Selected sample student responses to true/false statements.
- Appendix 5: Matched pretest and posttest data.

ACKNOWLEDGMENTS

We would like to thank Shannon Harris for statistical assistance during development. We would also like to acknowledge the support of the ASM education office including Amy Chang and Kelly Gull. Other faculty who have made this project possible by delivering the survey during development are Jamie Cunningham, Alfred Buchanan, Valerie Narey, Ann McDonald, Moria Nagy, M. Julia Massimelli, Sue Katz-Amburn, Brenden Mattingly, and Sonja Yung. This research was supported in part by a grant from the National Science Foundation, Division of Undergraduate Education (DUE), Conference Proposal I625772. The authors declare that there are no conflicts of interest.

REFERENCES

1. Hestenes D, Wells M, Swackhamer G. 1992. Force concept inventory. *Phys Teach* 30:141–151.
2. Klymkowsky MW, Garvin-Doxas K. 2008. Recognizing students' misconceptions through Ed's tools and the biology concept inventory. *PLOS Biol* 6:e3.
3. Garvin-Doxas K, Klymkowsky M, Elrod S. 2007. Building, using, and maximizing the impact of concept inventories in the biological sciences: report on a National Science Foundation-sponsored conference on the construction of concept inventories in the biological sciences. *Life Sci Educ* 6:277–282.
4. Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century. 2003. *BIO 2010: transforming undergraduate education for future research biologists*. The National Academies Press, Washington, DC.
5. Sadler PM. 1998. Psychometric models of student conceptions in science: reconciling qualitative studies and distracter-driven assessment instruments. *J Res Sci Teach* 35:265–296.
6. D'Avanzo C. 2008. Biology concept inventories: overview, status, and next steps. *Bioscience* 58:1079–1085.
7. Deane T, Nomme K, Jeffery E, Pollock C, Birol G. 2014. Development of the biological experimental design concept inventory (BEDCI). *CBE Life Sci Educ* 13:540–551.
8. Smith MK, Wood WB, Knight JK. 2008. The genetics concept assessment: a new concept inventory for gauging student understanding of genetics. *CBE Life Sci Educ* 7:422–430.

9. Marbach-Ad G, Briken V, El-Sayed NM, Frauwirth K, Fredericksen B, Hutcheson S, Gao LY, Joseph SW, Lee V, McIver KS, Mosser D, Booth Quimby B, Shields P, Song W, Stein DC, Yuan RT, Smith AC. 2009. Assessing student understanding of host-pathogen interactions using a concept inventory. *J Microbiol Biol Educ* 10:43–50.
10. Kalas P, O'Neill A, Pollock C, Birol G. 2013. Development of a meiosis concept inventory. *CBE Life Sci Educ* 12:655–664.
11. Couch BA, Wood WB, Knight JK. 2015. The molecular biology capstone assessment: a concept assessment for upper-division molecular biology students. *CBE Life Sci Educ* 14:ar10.
12. Stefanski KM, Gardner GE, Seipelt-Thiemann RL. 2016. Development of a lac operon concept inventory (LOCI). *CBE Life Sci Educ* 15:ar24.
13. Knudson D. 2006. Biomechanics concept inventory. *Percept Mot Skills* 103:81–82.
14. Merkel S and the ASM Task Force on Curriculum Guidelines for Undergraduate Microbiology. 2012. The development of curricular guidelines for introductory microbiology that focus on understanding. *J Microbiol Biol Educ* 13(1):32–38.
15. Horak R, Merkel S, Chang A. 2015. The ASM curriculum guidelines for undergraduate microbiology: a case study of the advocacy role of professional societies in reform efforts. *J Microbiol Biol Educ* 16(1):100–104.
16. Adams WK, Wieman CE. 2011. Development and validation of instruments to measure learning of expert-like thinking. *Int J Sci Educ* 33:1289–1312.
17. Creswell JW. 2013. *Qualitative inquiry & research design: choosing from among five approaches*, 3rd ed. Sage Publications, Thousand Oaks, CA.
18. Stemler SE. 2004. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Pract Assess Res Eval* 9(4):1.
19. Briggs AG, Hughes LE, Brennan RE, Buchner J, Horak REA, Amburn SK, McDonald AH, Primm TP, Smith AC, Stevens AS, Yung SB, Paustian TD. 2017. Concept inventory development reveals common student misconceptions about microbiology. *J Microbiol Biol Educ* 18(3):1–9.
20. Ericsson KA, Simon HA. 1998. How to study thinking in everyday life: contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind Culture Act* 5:178–186.
21. Ding L, Chabay R, Sherwood B, Beichner R. 2006. Evaluating an electricity and magnetism assessment tool: brief electricity and magnetism assessment. *Phys Rev Spec Top Phys Educ Res* 2: 010105-1-7.
22. Crocker L, Angina J. 1986. *Introduction to classical and modern test theory*. Harcourt Brace Jovanovich College Publishers, Orlando, FL.
23. Cunningham BC, Hoyer KM, Sparks D. 2015. Stats in brief: gender differences in science, technology, engineering and mathematics (STEM) interest, credits earned, and NAEP performance in the 12th Grade. National Center for Education Statistics, Institute of Education Sciences, US Department of Education, Washington, DC.
24. National Science Foundation, 2015. National Center for Science and Engineering Statistics. Women, minorities, and persons with disabilities in science and engineering: 2015. Special Report NSF 15-311.
25. Hake RR. 1998. Interactive-engagement vs. traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am J Phys* 66(1):64–74.
26. Aubrecht II GJ, Aubrecht JD. 1983. Constructing objective tests. *Am J Phys* 51(7):613–620.
27. Paustian TD, Briggs AG, Brennan RE, Boury NE, Buchner J, Harris S, Horak REA, Hughes LE, Katz-Amburn DS, Massimelli MJ, McDonald AH, Primm TP, Smith AC, Stevens AM, Yung SB. Development, validation, and application of the microbiology concept inventory. *J Microbiol Biol Educ* 18(3):1–10.