Paper ID #20545

Using Multiple Choice Responses to Assess Uncertainty in Student Understanding of Vector Concepts

Mr. Ramin Madarshahian, University of South Carolina

Ramin is a Ph.D. candidate in Structural engineering with two Master of Science degrees in Statistics and Structural Engineering. His focus of research is uncertainty quantification in engineering problems and model updating.

Mr. Chao Chen

Chao is a PhD student in the Department of Computer Science & Engineering at University of South Carolina. He is interested in applying machine learning algorithms and Bayesian statistics in social science study.

Dr. Juan M Caicedo, University of South Carolina

Dr. Caicedo is a Professor at the Department of Civil and Environmental Engineering at the University of South Carolina. His research interests are in structural dynamics, model updating, and engineering education. He received his B.S. in Civil Engineering from the Universidad del Valle in Colombia, South America, and his M.Sc. and D.Sc. from Washington University in St. Louis. Dr. Caicedo's educational interests include the development of critical thinking in undergraduate and graduate education. More information about Dr. Caicedo's research can be found online at http://sdii.ce.sc.edu

Dr. A. Steve McAnally, University of South Carolina

Steve McAnally is Associate Professor in the Department of Civil and Environmental Engineering at the University of South Carolina. His research interests focus on water and wastewater treatment, particularly appropriate technology applications for developing communities. Other interests include reform in undergraduate civil engineering education.

Dr. Charles E. Pierce, University of South Carolina

Dr. Pierce is the Director for Diversity and Inclusion and Associate Professor in the Department of Civil and Environmental Engineering at the University of South Carolina. He is a USC Connect Faculty Fellow for Integrative Learning, and a Bell South Teaching Fellow in the College of Engineering and Computing. Dr. Pierce also serves as the ASEE Campus Representative for USC.

Dr. Gabriel Terejanu, University of South Carolina

Gabriel Terejanu has been an Assistant Professor in the Department of Computer Science and Engineering at University of South Carolina since 2012. Previously he was a Postdoctoral Fellow at the Institute for Computational Engineering and Sciences at University of Texas at Austin. He holds Ph.D. in Computer Science and Engineering from University at Buffalo. He is currently working on the development of a comprehensive uncertainty quantification framework to accelerate the scientific discovering process and decision-making under uncertainty. Some projects currently supported by NSF and VP for Research include discovery of novel catalytic materials for biorefinery industry, modeling and prediction of naturally occurring carcinogenic toxins, and development of statistical models for tracking individual student knowledge.

Using Multiple Choice Response Options to Assess Uncertainty in Student Understanding of Vector Concepts

Introduction

Assessing conceptual understanding in large engineering courses is a challenging task. When we consider that assessment in engineering education is often performed in a deterministic fashion and does not include uncertainty, the challenge is even greater. Arguably, including uncertainty in student assessment could lead to a more meaningful measurement of conceptual knowledge (i.e., there is a 70% likelihood that a student knows a concept). Asking students to quantify the confidence in their answers would provide additional information towards the development of a framework that incorporates uncertainty in the assessment process. However, this requires training students in understanding how to provide a meaningful answer, something that might be difficult to do for most engineering courses.

Multiple-choice questions are commonplace in engineering assessments of student knowledge. Their prevalence is correlated with a number of factors that include student familiarity, ease of implementation, and simplicity of grading. Multiple-choice formats are used for formative assessment (e.g., in-class quizzes, which can be paper-based or embedded within classroom response systems) and summative assessment (e.g., midterm and final exams). Multiple-choice is often used for standardized tests, as well, such as the Fundamentals of Engineering exam. Given their widespread use, there are ample opportunities for integrating assessments of student uncertainties into multiple-choice questions, which can be particularly useful in large engineering classes. Of course, the consideration of student uncertainties, self-confidence, partial knowledge, and guessing on multiple choice tests is not new and has been studied in the educational literature (e.g., Burton 2005; Bereby-Meyer et al. 2003; Burton 2001; Burton & Miller 1999; Ben-Simon et al. 1997; Hassmén & Hunt 1994).

In this work, we developed a modified approach for student responses to multiple-choice questions using paper-based quizzes on concepts associated with vectors. This approach is called the coin distribution response method, which was coupled with the traditional approach of choosing a single answer. This paper presents our findings from implementation trials in two semesters of a statics course in civil and environmental engineering.

Methodology

Coin Distribution Response Method

The coin distribution response method, or CDRM, was tested as part of a larger project to create a Bayesian network model of student conceptual understanding of vectors (Chen et al. 2016). For each multiple-choice question, students are provided with a hypothetical allotment of 100 coins that must be distributed across the answer choices. The students are informed that the coin assignments should represent the likelihood that each answer is the correct one. Thus, an assignment of 100 coins to a single answer choice represents complete confidence that it is the correct answer. A number of coins less than 100 equates to a lower degree of confidence. On the other hand, an assignment of 0 coins to an answer choice represents complete confidence that it is not the correct answer. The distribution of coins on each question can provide a rich source of information regarding the individual student's perceived level of understanding.

The CDRM was implemented in back-to-back semesters, fall 2015 and spring 2016, in a statics course. Three mini-quizzes were administered within a three-week span; each mini-quiz was designed to be short, with a maximum of four multiple-choice questions. Students completed each quiz using the traditional method of selecting one answer per question, and quizzes were graded in a deterministic fashion based on right or wrong answers. Students were also required to use CDRM in conjunction with their answer choice, even though the coin distributions had no effect on their scores.

Questions for each mini-quiz were created to address conceptual understanding of vector representation of properties and vector summation in two and three dimensions. Future work will consider expansion and integration of CDRM with other concepts using the Statics Concept Inventory (e.g., Steif and Dantzler 2005). While the research context is based in student understanding of vector concepts in statics, this paper does not elaborate on proposed correlations between the multiple choice response data and predicted level of knowledge of vectors. That work is in progress. Rather, this paper concentrates on the student response patterns when using CDRM for the first time in a statics course.

Implementation Trials in Statics

Fall 2015. Course enrollment in the first semester of this implementation trial was 37 students. Fall is considered the “off” semester for statics, which means that students who progress through the curriculum in a traditional sequence do not take it in the fall. There tends to be a smaller enrollment in the fall as compared to the spring. Three short quizzes were distributed: quiz 1 was comprised of 2 questions; quiz 2 was effectively comprised of 1 question; and quiz 3 was comprised of 3 questions. The maximum number of student answers, or counts, is 222 based on a total of 6 questions in a class of 37 students. Due to student absences, the actual count was 171.

The second quiz was designed with a maximum of 4 questions. The first question was required of all students; one or more of the remaining 3 questions were assigned to a small group of students. These assignments were based on their selection of incorrect answers to questions on quiz 1. However, this population was small, yielding just 5 additional counts. Student absences for quiz 1 and/or 2 contributed to the low number of counts. In other words, one of the challenges with personalizing quizzes for students is that those students who are struggling with conceptual understanding might also not be attending class. In the spring semester, it was decided to assign all 4 questions on quiz 2 to all students, regardless of performance on quiz 1.

Spring 2016. Course enrollment in the second semester of this implementation trial was 55 students. The same quiz questions were used; as noted above, quiz 2 was implemented with all 4 questions. The maximum number of counts is 495, based on a total of 9 questions in a class of 55 students. Due to student absences, the actual count was 446, which is equivalent to a 90% completion rate. This is higher than the completion rate of 77% in fall 2015.

Results

One of the research questions was to understand how students would assign coins to the correct answer as a function of whether or not it was selected as the correct answer. In other words, if a student selects the correct answer, how does that student assign the likelihood that this answer is in fact the correct answer, based on the number of coins? On the other hand, if a student selects an incorrect answer, how does that student assign the likelihood that another answer could, in fact, be the correct answer even though it is not chosen? Theoretically, this can be used as a relative measure of self-belief in knowledge compared to accurate knowledge of a concept. The cumulative results of coin assignments across all of the multiple-choice questions are thus presented for those selecting correct answers and those selecting incorrect answers.

Selections of Correct Answers

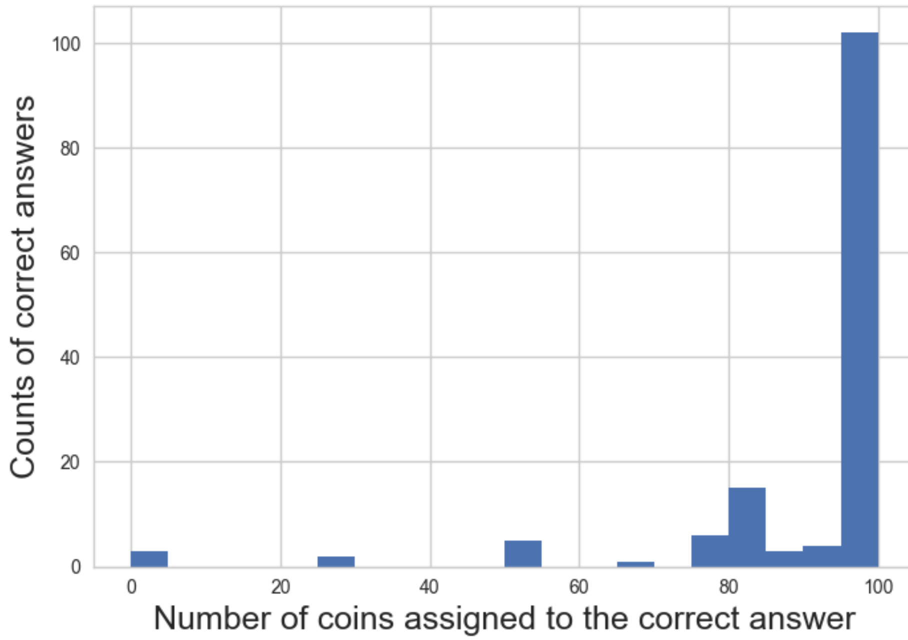
Figure 1 illustrates the student counts of coins assigned to the correct answer when selecting the correct answer. Results are split into fall 2015 and spring 2016 semesters; however, the trends in student responses are similar. First, it is clear that students selected the correct answers at a high rate, greater than 70% in both semesters. Thus, there is significantly more available data for interpreting student self-belief in knowledge when their knowledge leads to choosing the correct answer. Student self-confidence in knowing the correct answer is also quite high, given that more than 75% of students assigned 96-100 coins. There is a much smaller population of students, on the order of 10-15%, who assigned less than 80 coins to the correct answer. Although this is a somewhat arbitrary threshold, an assignment of less than 80 coins might suggest that these students have some reasonable degree of uncertainty in knowing the correct answer.

Selections of Incorrect Answers

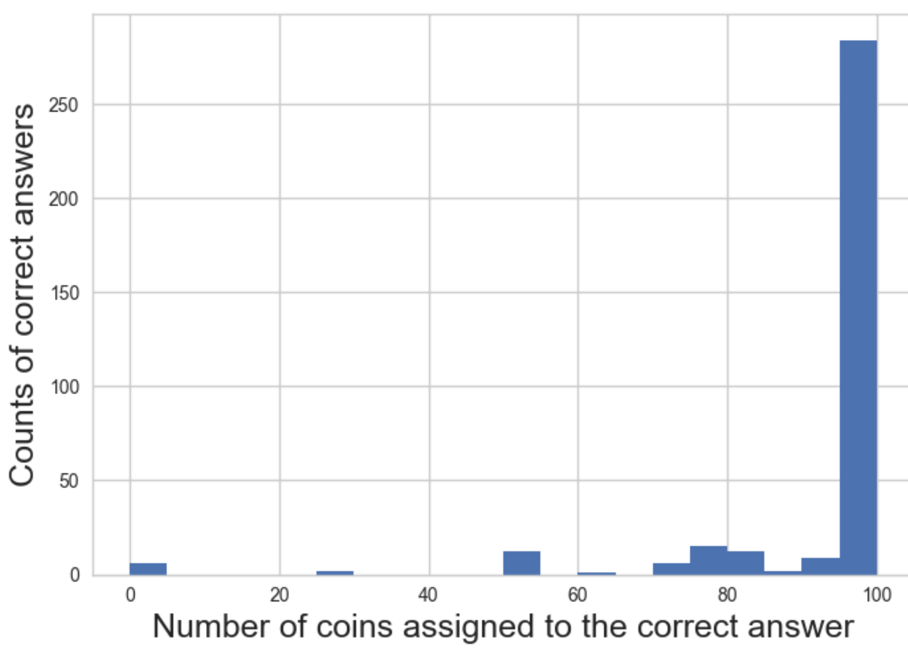
Figure 2 illustrates the student counts of coins assigned to the correct answer when selecting the incorrect answer. As noted above, there is much less data to evaluate this condition. With that said, the trends in student responses are consistent between the two semesters. At least half of the students assigned 0-5 coins to the correct answer. This response behavior represents either a deep-rooted misconception or misunderstanding, or an unwillingness to use the coin distribution method as a reflection of one's uncertainty in knowing the answer. In either case, these student responses do not provide much more information than the traditional, deterministic approach for answering multiple-choice questions. There is, however, a small population of students that provide some reasonable likelihood (>25 coins) that the correct answer could be correct, even though it was not the final selection. A threshold of 25 coins is used because that number represents random selection of an answer. In other words, a student with no knowledge would theoretically assign 25 coins to each one of four answer choices.

Coin Distribution Frequencies

The results shown in Figures 1 and 2 represent one measure of how the CDRM was utilized, in terms of the magnitude of coins (0 to 100) assigned to the correct answers. Figure 3 illustrates a second measure of utilization, in terms of the number of answer choices to which students

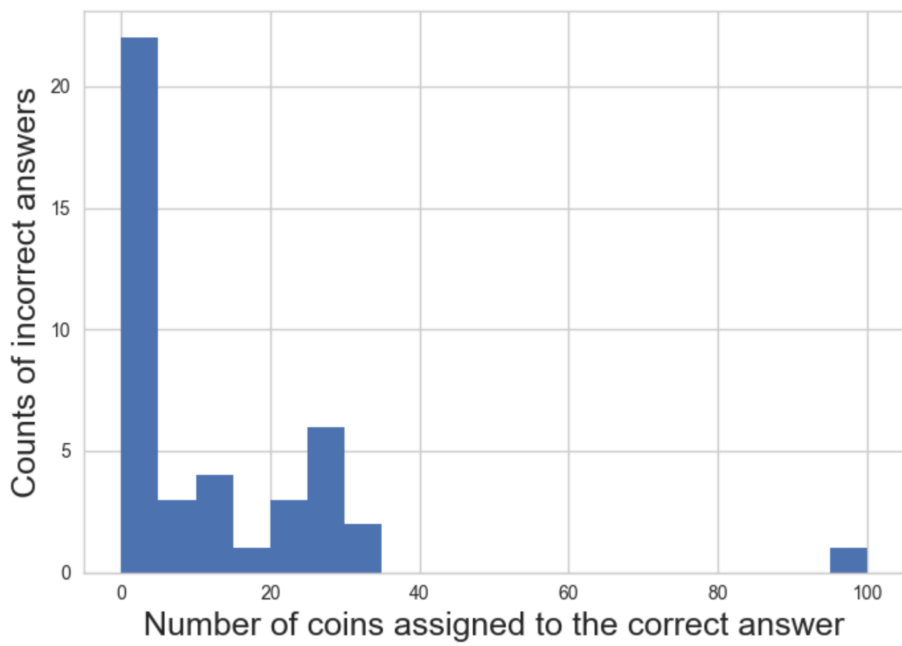


(a)

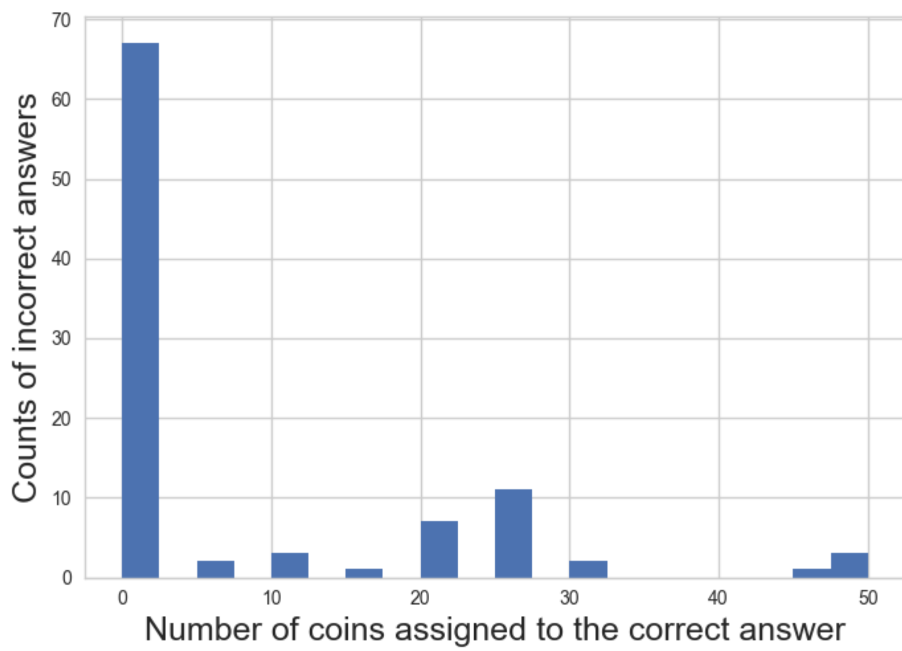


(b)

Figure 1. Number of coins assigned to correct answer when choosing *correct* answer for (a) fall 2015 and (b) spring 2016.

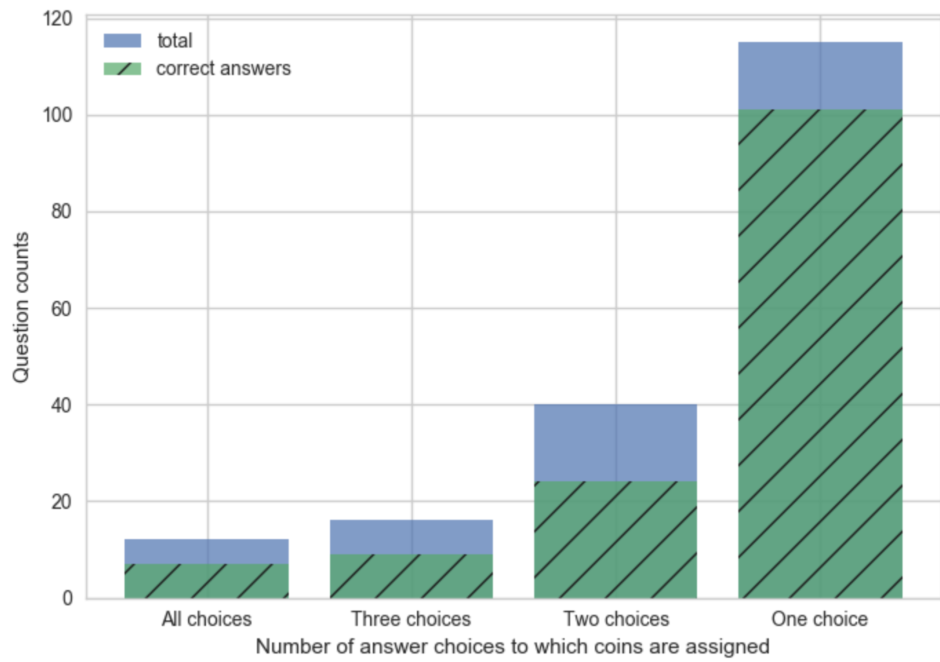


(a)

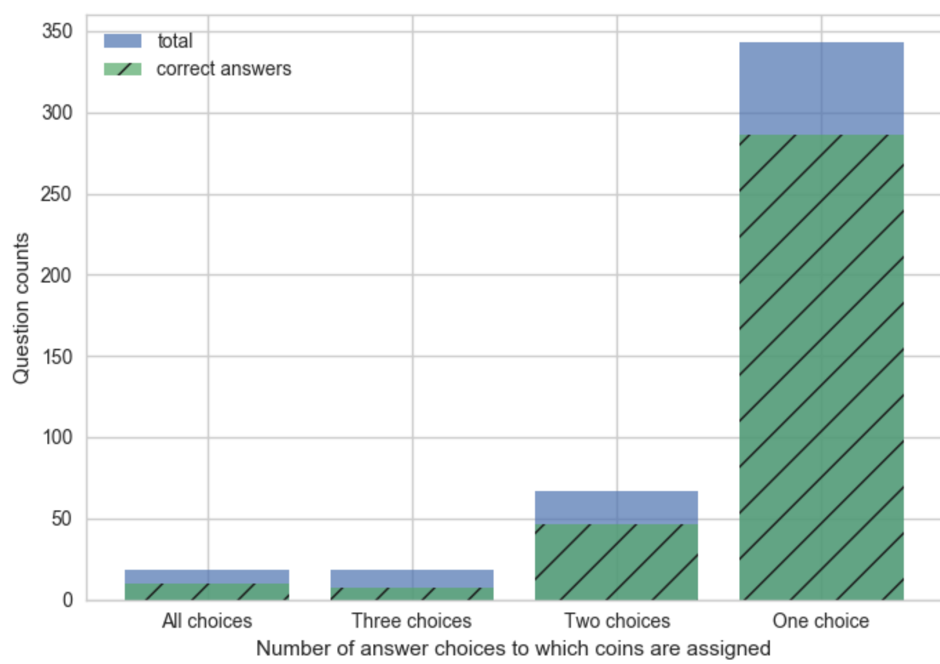


(b)

Figure 2. Number of coins assigned to correct answer when choosing *incorrect* answer for (a) fall 2015 and (b) spring 2016.



(a)



(b)

Figure 3. Cumulative counts of the number of answer choices that are assigned >0 coins for (a) fall 2015 and (b) spring 2016.

assigned a non-zero allotment of coins. The data show that students tend to place all of their coins on a single answer. In the fall 2015 semester, this response behavior represented about 67% of all question counts. It was even higher in spring 2016, with more than 75% of question counts having 100 coins assigned to one answer. When students choose to distribute coins across multiple choices, the data also show that two answer choices are more frequent than three or four choices.

Perhaps the most interesting results in Figure 3 are the frequencies of correct answers when distributing and not distributing coins. Students that do not distribute coins, meaning that 100 coins are assigned to one choice, are more likely to select the correct answer than those students who distribute coins across two or more choices. The patterns are remarkably consistent between the fall 2015 and spring 2016 semesters. In the fall and spring, respectively, 88% and 83% of the answers assigned 100 coins were correct. In comparison, 62% and 64% of the answers that students selected were correct when coins were distributed rather than assigning 100 coins to the selected answer. This means that answers were wrong more than one-third of the time when students expressed some uncertainties through the allocation of coins to multiple choices.

Discussion and Recommendations

Results show that students develop and use different patterns to assign coins. These patterns can change per question and per student. Some students tended not to distribute coins at all, meaning that 100 coins were placed on their answer for all quiz questions. Other students often distributed coins over several choices. In the latter case, the goal is to correlate this perceived degree of uncertainty with how well students know or do not know the specific concepts. This can be elucidated using student responses to specific questions. One example is with the first question on quiz 1. In spring 2016, eight students selected the same incorrect choice. However, the coin distribution patterns were different. Two of them split the 100 coins between this apparent distractor and the correct answer; two others distributed coins between the distractor and a second incorrect choice; and the remaining four students assigned all 100 coins to the distractor. An instructor using traditional multiple choice response methods would not be able to differentiate these student responses; in other words, each student would receive the same score of zero for that question. Yet, the coin distribution patterns might suggest that students who were unsure between the correct answer and the distractor might have a better understanding of the concept than those who assigned all 100 coins to the distractor.

There is also evidence of student mistakes when using the coin distribution response method. Figure 1, for example, shows that a small number of correct answers were selected but assigned 0 coins. This unexpected response does not make sense, and in some cases, it appears to be the result of a one-time error. Three different students made this mistake once. However, another student made this mistake six times, suggesting that he/she was unclear on how to use CDRM in conjunction with the selected answer.

There are two important observations from these implementation trials. First, a significant percentage of student answers to quiz questions did not have distributed coins, resulting in the assignment of 100 coins to the selected answer. In other words, it appears that a large population *might not* have applied a self-assessment of their confidence in the selected answer. Rather, these students might have relied on their pre-established behaviors to select a singular answer and overlaid that answer with an assignment of 100 coins. This is a justifiable approach from a student perspective, given that the coin distributions did not affect how students were graded. For future work, it is recommended that CDRM be integrated into the scoring of multiple-choice questions.

Second, the use of CDRM should be expanded to other assessment instruments. In this case, it might be better to evaluate this method on exams rather than short quizzes to ensure a higher rate of participation across the spectrum of student levels of conceptual understanding. Given that a fraction of students were absent for each quiz in each semester, it is hypothesized that some useful data on coin distributions were not acquired.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1504728.

References

Ben-Simon, A., Budescu, D.V., & Nevo, B. (1997). A comparative study of measures of partial knowledge in multiple-choice tests. *Applied Psychological Measurement*, 21(1), 65-88.

Bereby-Meyer, Y., Meyer, J., & Budescu, D.V. (2003). Decision making under internal uncertainty: the case of multiple-choice tests with different scoring rules. *Acta Psychologica*, 112(2), 207-220.

Burton, R.F. (2001). Quantifying the effects of chance in multiple choice and true/false tests: question selection and guessing of answers. *Assessment & Evaluation in Higher Education*, 26(1), 41-50.

Burton, R.F. (2005). Multiple‐choice and true/false tests: myths and misapprehensions. *Assessment & Evaluation in Higher Education*, 30(1), 65-72.

Burton, R.F., & Miller, D.J. (1999). Statistical modelling of multiple‐choice and true/false tests: ways of considering, and of reducing, the uncertainties attributable to guessing. *Assessment & Evaluation in Higher Education*, 24(4), 399-411.

Chen, C., Madarshahian, R., Caicedo, J.M., Pierce, C.E., & Terejanu, G.A. (2016). Bayesian network models for student knowledge tracking in large classes. ASEE Annual Conference & Exposition.

Hassmén, P., & Hunt, D.P. (1994). Human self‐assessment in multiple‐choice testing. Journal of *Educational Measurement*, 31(2), 149-160.

Steif, P.S., & Dantzler, J.A. (2005). A statics concept inventory: Development and psychometric analysis. *Journal of Engineering Education*, 94(4), 363.