An Evaluation of Student Responses to a Fluid Mechanics Concept Inventory

Ozge Uyanik, University of South Florida

Ozge Uyanik is a PhD candidate in the Department of Mechanical Engineering at the University of South Florida, where she researches engineering education and develops microfluidic systems and surface acoustic wave-based devices using 3D printing and advanced microfabrication techniques. She holds a B.S. from Istanbul Technical University (Turkey), having completed a double major in mechanical engineering and meteorological engineering. Ozge was awarded the 2024 Student Research Award by USF Research & Innovation and the Florida High Tech Corridor for her exceptional research accomplishments. This honor highlights her significant contributions to advancing the high-tech industry and fostering innovation. Since 2021, Ozge has served as a teaching assistant for various undergraduate and graduate courses. Additionally, she is engaged in NSF-supported research and scholarship focused on metacognition support and assessment within a flipped and blended Fluid Mechanics course. Ozge is affiliated with the Tau Beta Pi Engineering Honor Society and has held roles as professional chair and sub-lead Engineering Expo chair in the Florida Gamma Chapter. She is also an active member of the Society of Women Engineers (SWE) and currently co-chairs the graduate section of SWE at the University of South Florida. She has judged the SWE Recognition Program, emphasizing Diversity, Equity, Inclusion & Belonging. Recognizing her leadership potential, she has been selected for SWE's Collegiate Leadership Institute (CLI) Program and the SWE Leadership Development Program (LDP).

Prof. Autar Kaw, University of South Florida

Autar Kaw is a mechanical engineering professor at the University of South Florida. Recognized with the 2012 U.S. Professor of the Year Award (doctoral and research universities) from the Council for Advancement and Support of Education and the Carnegie Foundation for the Advancement of Teaching, Kaw's scholarly pursuits include engineering education research, adaptive, blended, and flipped learning, open courseware development, composite materials mechanics, and examining the future of higher education. His research has received funding from the National Science Foundation, Air Force Office of Scientific Research, Florida Department of Transportation, and Wright Patterson Air Force Base. Supported by the National Science Foundation, Kaw has led a national collaboration to develop, implement, refine, and assess online resources for open courseware in Numerical Methods (http://nm.MathForCollege.com). These resources gather over 1 million page views annually and 1.6 million YouTube lecture views, attracting more than 90,000 visitors to the "numerical methods guy" blog. This work also evaluates how flipped, blended, and adaptive environments effectively help engineering students learn content, hone group-working skills, and perceive their learning contexts. Kaw has published more than 130 peer-reviewed technical papers, and his opinion pieces have been featured in the Tampa Bay Times, the Tampa Tribune, and the Chronicle of Higher Education.

Dr. Rasim Guldiken, Oklahoma State University

My name is Rasim Guldiken. I am a John Brammer Endowed Professor and Head of the Mechanical and Aerospace Engineering Department at Oklahoma State University. I was also a Professor and the Associate Dean for Academic Affairs of the College of Engineering at the University of South Florida until 2025. I have taught fluid mechanics, thermodynamics, and engineering mathematics-related courses 40+ times to 3000+ B.S., M.S., and Ph.D. students for the last 17 years. My engineering education interests lie in open courseware for courses, metacognitive activities, and flipped learning. Dr. Guldiken has been recognized with multiple awards and honors, including the 2022 ASME Fellow, 2022 USF Faculty Outstanding Research Achievement Award, 2022 USF Academic Excellence Award, 2019 and 2012 USF University-Wide Outstanding Undergraduate Teaching Awards, 2018 USF Outstanding Graduate Faculty Mentor Honorable Mention and 2014 Society of Automotive Engineers (SAE) Ralph Teetor Educational Award.

An Evaluation of Student Responses to a Fluid Mechanics Concept Inventory

Abstract

A concept inventory is a valuable tool to assess a student's grasp of specific concepts within a course. In this study, we utilized the Fluid Mechanics Concept Inventory (FMCI) in a mandatory Fluid Mechanics course at a large southeastern university offering an undergraduate Mechanical Engineering program. The FMCI comprised 30 multiple-choice questions, each with three to five options. To gather data, we administered the inventory on campus for 60 minutes in a 75-minute class period during the penultimate week of the semester, with 216 participating students across the Fall 2022, Spring 2023, and Fall 2023 semesters. Our analysis focused on various parameters, including item difficulty, item discrimination, reliability, alpha with items deleted, item response theory, and subscale alpha. The calculated Cronbach's alpha as a measure of reliability was found to be 0.73, and the correlation between the concept inventory and the final exam score was weak. We identified several questions that appeared to warrant removal based on different criteria. Seven questions were suggested for removal when considering the combination of accepted ranges of difficulty (0.3 to 0.9) and discrimination indices (≥0.2), while according to the item-response-theory-analysis criterion, only four questions warranted removal. To ensure comparability with other studies using the same inventory, we used all questions despite the identification of potential removal candidates. We could not perform exploratory or confirmatory factor analysis or diagnostic classification modeling due to the design of the questions, which addressed distinct concepts by fewer than three questions converging on a single concept. This limitation highlights the need for a revised concept inventory to enhance student learning in this area and ensure the effectiveness of our teaching methods. The paper outlines a step-by-step Delphi method for developing the revised FMCI.

1. Introduction

The aim of engineering courses, similar to other fields, is to enhance students' comprehension of basic concepts [1]. Fundamental concepts are not merely lists of topics that make up the curriculum but the core ideas of a subject that form the basis for equations and problem-solving [2]. If students have not developed a good understanding of essential concepts, they will have an unclear foundation of knowledge. This unclarity may result in students being unable to effectively apply course concepts to solve real-world problems not explicitly addressed in class, leaving them unprepared for advanced study [3].

Engineering problems require conceptual and procedural knowledge [4]. Instructors are increasingly broadening learning objectives in undergraduate STEM courses to include not only procedural knowledge but also conceptual understanding [5]. Conceptual knowledge refers to principles and relationships within a domain. Procedural knowledge, also known as skills, algorithms, or strategies, consists of sequences of actions to solve problems [6]. When conceptualizing undergraduate education, emphasis is placed on how students' ideas fit into normative scientific and engineering explanations and practices, and how ideas that do not fit will be changed. As with all learning, the undergraduate science and engineering learning experience builds on students' prior knowledge [7]. Developing and assessing students' computational skills can be accomplished easily through traditional homework assignments, lectures, and standardized exams. In contrast, developing and evaluating conceptual understanding is more complex [8]. To address this challenge, concept inventories (CI) have been designed to assess students' conceptual understanding [5].

In this paper, an existing Fluid Mechanics Concept Inventory (FMCI) [9, 10] was administered to students of a Fluid Mechanics course in the Department of Mechanical Engineering at a large southeastern USA university. The FMCI consists of 30 multiple-choice questions. It was given over 60 minutes during a 75-minute class period in the penultimate week of the semester to 232 students from the Fall 2022, Spring 2023, and Fall 2023 semesters. Data from 216 students who chose to participate were used in the study and analyzed to evaluate the reliability and quality of the questions. The analysis identified questions that may need to be removed. Since numerous concepts are being tested or sub-grouped in the current FMCI, it is recommended to be revised using the Delphi technique [11, 12] to test on major concepts only. Only those questions from the current FMCI that pass the item analysis and fit the chosen concepts will be retained. This approach would ensure that the revised FMCI accurately measures students' understanding of the specific concepts of the course.

2. General Concept Inventory Related Literature

Concept inventory, an assessment method focused on testing understanding of the fundamental concepts within a subject, avoids assessing procedural knowledge. Therefore, CI questions do not usually involve calculations or procedural algorithms [5, 13]. CIs include multiple independent concepts in a single test format [14], usually consisting of multiple-choice questions designed to test students' grasp of core subject concepts and identify possible misconceptions [13]. The answer choices comprise a correct option and several incorrect alternatives known as distractors [15]. When designing these inventories, distractors are selected to be particularly appealing to students with common misunderstandings of the material [3]. CIs help identify common errors students make during tests and provide valuable information for developing teaching strategies [14]. Because concept inventories reveal the difference between the intended learning outcomes and what students learn, educators aiming to enhance their instructional effectiveness can utilize CIs as a benchmark to measure progress in bridging this gap [16].

The work led by David Hestenes [17, 18] has contributed significantly to the widespread use of CI as an assessment tool across engineering and various other disciplines [15, 16]. The Mechanical Diagnostic Test was designed and validated to evaluate students' basic knowledge in Introductory Physics courses. Initial versions required written responses; however, the final version adopted a multiple-choice format, incorporating answers that reflected prevalent misconceptions as alternatives [16, 18]. To enhance the Mechanics Diagnostic Test, the Force Concept Inventory (FCI) was created to evaluate students' general understanding of the Newtonian concept of force. The inventory is an investigation of belief systems rather than an intelligence test, with errors in the inventory providing more insight than correct choices [17]. The Force Concept Inventory (FCI) is regarded as a thorough gauge of students' understanding of Newtonian physics. The variations in pre- and post-test results within a class can be utilized to assess the impact of various instructional methods on student achievement. Consequently, the FCI has had a catalytic impact by promoting discussion of optimal teaching practices [19].

Successful implementation of the FCI has contributed to the creation of concept inventories in various areas of engineering, science, and mathematics [13, 20, 21]. The Foundation Coalition (FC), funded by NSF, began efforts to develop Concept Inventories across engineering disciplines in 2000 [22]. FC provides concept inventories for various fields, including Circuits, Dynamics, Chemistry, Computer Engineering, Fluid Mechanics, Materials, Electromagnetics, Heat Transfer, Electronics, Strength of Materials, Signals and Systems, Thermodynamics, and Waves [23]. With NSF funding, a Colorado School of Mines research team developed the Thermal and Transport Concept Inventory. The inventory covers introductory thermodynamics, fluid mechanics, and heat transfer concepts [12, 16, 24, 25]. In addition, the Statistics Concept Inventory was created to assess students' understanding of basic statistics concepts [26, 27]. A similar goal was pursued by developing the Statics Concept Inventory for fundamental concepts of statics [28]. Kaw et al. [29] developed a concept inventory for an engineering course in Numerical Methods starting in 2013. They used the Delphi technique [11, 12] to identify six key concepts and refined the concept inventory over a decade to an 18-question test.

3. Fluid Mechanics Concept Inventory Related Literature

Fluid mechanics, one of the fundamental core courses in undergraduate engineering education [30], includes the mechanics and dynamics of fluids. It is based on fundamental physics and Newtonian mechanics [31]. In 2001, faculty members from the Mechanical Engineering departments at the University of Wisconsin-Madison and the University of Illinois, Urbana-Champaign, initiated a collaboration to develop a Fluid Mechanics Concept Inventory (FMCI) [9, 31]. This inventory was designed explicitly for the Department of Mechanical Engineering undergraduate Fluid Mechanics course.

A comprehensive process involving faculty input was utilized to identify the essential concepts in developing the FMCI. The faculty-developed list of approximately 25 distinct concepts was categorized into three main areas: basic concepts, fundamental fluid relations, and special cases. To further validate these concepts, students were requested to compose a list of 10 concepts they felt confident about and deemed necessary, and 10 concepts they felt uncertain about and deemed unimportant. The student feedback facilitated refining the list of key topics, leading to the development of a set of questions[10, 32]. The objective was to ascertain whether students grasped the fundamental concepts without performing any calculations. Consequently, numerical calculations were omitted in alignment with the structure of the FCI. The FMCI comprised 30 multiple-choice questions, including graphical and visual representations of the concepts under study [10, 30, 33].

A study by Fraser et al. [34] used a reduced version of the FMCI to examine the effect of computer simulations on engineering students' comprehension of fluid mechanics. The questions not considered integral to the course were mostly about boundary layers. The reduced version of the test was administered mid-semester to students enrolled in a sophomore fluid mechanics course. Analysis of the test results by question category highlighted three principal areas of difficulty: pressure measurement, variable diameter pipe flow, and velocity profiles between flat plates. Therefore, three sets of simulations were developed to address these identified areas. Comparison of the post-test (using the same reduced FMCI) with the pre-test results allowed the effectiveness of the simulations to be evaluated [34].

Baghdanov [35] investigated students' comprehension of pressure pipeline flow. The study involved participants with distinct experience levels regarding fluid mechanics who were interviewed using questions from the FMCI and open-ended queries concerning pressurized pipe systems. The findings revealed that while students could readily grasp the changes in velocity within pressure pipelines of varying diameters, they often struggled to predict the corresponding pressure changes correctly. On the other hand, a common misconception related to horizontal pipelines was that water is considered compressible. Additionally, participants' approaches to vertical pipelines included using hydrostatic pressure to determine the change in pressure [35].

Watson et al. [32] sought to undertake the initial steps necessary to develop a Fluid Mechanics concept inventory for civil engineering majors by modifying the FMCI designed for mechanical engineering. Key topics relevant to civil engineers were identified through input from an expert panel, student feedback, and analysis of the junior civil engineering students' performance on the current concept inventory. There was general agreement on fundamental fluid properties, conservation of mass, Bernoulli's equation, and conservation of momentum. On the other hand, compressible flow and boundary effects were considered out of scope. Additional concepts deemed necessary for inclusion in a prospective fluid mechanics concept inventory in civil engineering were also discussed.

In another study [30], a flipped classroom approach used in an undergraduate-level Fluid Mechanics course was compared with a concurrent section presented in a traditional mode. Student success was evaluated by comparing pre-test and post-test FMCI Exam scores, average final exam scores, and multivariate regression analysis. The FMCI was selected since its conceptual structure enabled pre- and post-testing opportunities to compare results.

In the study conducted by Brown and Barner [36], responses were gathered from approximately 100 civil engineers regarding concept inventories for Statics, Strength of Materials, and Fluid Mechanics. The general scores and response patterns of each concept inventory question were compared between students and engineers. The average student score for the Fluid Mechanics concept inventory was 50.3%, and they outperformed engineers in 12 out of 30 questions. On the other hand, engineers with an average of 11 years of experience scored 40.6% and outperformed students on only one question. Moreover, interviews consisting of eight questions from the Fluid Mechanics concept inventory were conducted with civil engineering professionals and students.

A French version of FMCI was also created with minor modifications, grouping the questions into categories: statics of fluids, ideal fluids and conservation laws, external viscous flows, and internal viscous flows [33].

4. Statistical Analysis of Fluid Mechanics Concept Inventory

According to Jorion et al. [5], using a concept inventory (CI) involves three general claims. First, students' overall performance can demonstrate their understanding of the CI concepts. Second, multiple questions on each topic allow for assessment of students' grasp of specific concepts.

Third, it can reveal students' vulnerability to misconceptions or common mistakes, as evidenced by their answers. In this study, we use the classical test theory (CTT) and item response theory (ITR) to analyze the Fluid Mechanics Concept Inventory (FMCI) to examine these claims. Jorion et al. [5] also used exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and diagnostic classification modeling (DCM) in their analytical framework; we did not use these, as the FMCI concept classification is either unclear or broad.

4.1 Questions by Concepts

The FMCI consists of questions on several concepts, including introductory fluid statistics, Bernoulli's equation, conservation principles, drag force, and boundary layer. Watson et al. ([32] Table 1) show that at least 16 concepts are covered in the 30 questions. In their paper, each question is marked by a corresponding targeting concept.

The course instructor, the paper's last author, classified the questions into broader concept areas, which are given in Table 1. Although prior knowledge concepts may be derived from many concepts, we bundled them as one - Introductory Concepts. The number of questions asked by each concept is given in Table 1. As one can note, there are six concepts, each with three or more questions.

Table 1. Questions by concepts

	Concept	Questions
1	Introductory Concepts	5, 7, 15, 16, 32
2	Fluid Statics	9, 10, 12, 14, 18, 30
3	Bernoulli's equation	4, 11, 17, 23, 27
4	Conservation Principles	3, 6, 13, 22, 25, 31
5	Drag force	19, 28, 29
6	Boundary Layer	8, 20, 21, 24, 26

4.2 Participation

Table 2 shows the number of students enrolled and participating in the study. The study was part of an NSF grant [37], and the existing FMCI was one of the assessment tools used to compare a flipped Fluids Mechanics course with and without using metacognition exercises [38, 39]. A total of 232 were enrolled in the three semesters, while 216 participated.

Table 2. Number of enrolled and study participants in each semester

Semester	Number of Enrolled Students	Number of Study Participants
Fall 2022	102	97
Spring 2023	66	58
Fall 2023	64	61
Total	**232**	**216**

4.3 Classical Test Theory Results

Table 3 shows each question's classical test theory results and the averages and ranges of the numbers calculated for the FMCI. Each column in the table is described below. These results were obtained using open-source software [40] that analyzes multiple-choice questions. The web-based software is user-friendly and needs only student responses and an answer key in CSV format.

Cronbach Alpha without item (CAWI): These alphas measure test reliability by excluding a question. If the reliability is higher by excluding the question, that is, CAWI is higher than the overall Cronbach Alpha, the question has negatively impacted the test's reliability. That question may become an item for removal.

The Difficulty Index (DIFFI): This index gauges the percentage of students who answered a test item correctly. The difficulty index ranges from 0 to 1, where, nonintuitively, values closer to 1 indicate easier items (a higher number of students answered correctly), and values closer to 0 signify more challenging items (a higher number of students answered incorrectly). A low difficulty index is not preferred. For instance, if a question has four choices, a difficulty index below 0.25 can be achieved by random guessing. Similarly, a high difficulty index (greater than 0.9) is also not favored as it does not differentiate between low- and high-scoring students.

Discrimination Index (DISCI): This index evaluates an item's ability to distinguish between high- and low-scoring students. Positive values suggest that students who performed well on the overall test answered this question correctly, while those who did poorly were inclined to answer it incorrectly. Conversely, negative values imply that lower-scoring students were more likely to answer the question correctly, whereas higher-scoring students tended to err, indicating a need for item review. Values close to zero imply that the item does not distinguish between high- and low-performing students. The index is calculated using the scores of the top 27% and bottom 27% of the class on an exam [41]. It is determined by subtracting the number of correct responses of the low group from those in the high group and then dividing by the total class size. The index ranges from -1.0 to +1.0, where values approaching 0.0 denote no discrimination. An ideal index approaches 1.0, with a value of 0.3 or above, which is considered highly discriminating.

The Point-Biserial Correlation Coefficient (PBCC): This coefficient measures the Pearson correlation between a dichotomous variable (such as a dichotomously scored item: correct/incorrect) and a continuous variable (like the overall test score). It is considered a simpler method to determine the discrimination index. The coefficient ranges from -1.0 to +1.0, with values close to 0.0 indicating no discrimination between high and low performers. An index close to 1.0 is ideal, and any value of 0.3 or higher signifies strong discrimination.

Modified Point-Biserial Correlation Coefficient (MPBCC): This statistic is the same as PBCC, but item scores are correlated to overall test scores without considering the given item in the overall test score. It is stricter than the PBCC and can be viewed the same way.

Number of options: This refers to the count of options in the multiple-choice test question. The FMCI has a varied number of options across its questions.

Number of Poor Distractors (NPD): An accepted rule of thumb is that at least 5% of respondents should select distractors for a question [42]. This column in Table 3 shows the number of poor distractors in a question that fell below the 5% threshold. If the FMCI is to be revised, these distractors must be reconsidered.

The reliability Cronbach alpha for the whole test is 0.731, with an average difficulty index of 0.524 and a discrimination index of 0.345. Table 3 also gives the ranges for the various statistics. The following section explains how these statistics are used to evaluate individual questions.

4.4 Candidate Questions for Removal from FMCI

In a typical multiple-choice test, the statistical measures in Table 3 are used to decide which questions should be removed and which should be kept. These measures may also indicate whether to review questionable items. The four criteria we used based on the statistical measures of Table 4 are described below.

1) Jorion et al. Criterion [5] is commonly used for concept inventories. Some examples include Statics, Statistics, and Dynamics CIs [26, 28, 43]. Acceptable questions have a Difficulty Index between 0.3 and 0.9, and the Discrimination Index is greater than 0.2. We mention other criteria below so the reader can judge and apply them to other multiple-choice question examinations they may provide.
2) AlphaWOI Criterion is where we include acceptable questions when the test's reliability is higher than when that question is removed.
3) The Stringent Criterion is more severe than the Jorion et al. criterion [5]. Acceptable questions have a Difficulty Index between 0.3 and 0.9, and the Point-Biserial Correlation Coefficient is more than 0.3.

Table 3: Statistics measures for each question

Question	Title	Concept	Alpha WOI	Difficulty	Discrimination	PBCC	MPBCC	Number of options	Number of Poor distracto
Q3	Continuity; compressible	4	0.725	0.664	0.387	0.334	0.239	5	2
Q4	Bernoulli; incompressible	1	0.712	0.749	0.562	0.528	0.456	4	1
Q5	Boundary conditions	3	0.716	0.370	0.486	0.461	0.373	4	0
Q6	Momentum; incompressible	4	0.714	0.379	0.559	0.493	0.408	5	0
Q7	Pressure definition	1	0.729	0.915	0.125	0.209	0.150	4	2
Q8	Boundary layers; incompressible	6	0.723	0.507	0.480	0.365	0.266	5	1
Q9	Pascal's Law	2	0.725	0.810	0.279	0.313	0.234	5	3
Q10	Manometry; compressible	2	0.719	0.341	0.411	0.414	0.324	5	0
Q11	Bernoulli; incompressible	3	0.722	0.483	0.406	0.379	0.281	3	0
Q12	Forces on submerged surface	2	0.717	0.583	0.469	0.443	0.352	5	1
Q13	Ideal Gas Law	4	0.726	0.327	0.382	0.317	0.222	5	0
Q14	Manometry; compressible	2	0.725	0.341	0.351	0.332	0.237	5	1
Q15	Shear stress; compressible	1	0.722	0.374	0.387	0.380	0.286	5	2
Q16	Boundary layers; incompressible	1	0.744	0.592	0.034	0.047	-0.058	4	1
Q17	Bernoulli; incompressible	3	0.715	0.706	0.517	0.482	0.401	4	1
Q18	Manometry; compressible	2	0.725	0.403	0.416	0.339	0.241	5	1
Q19	Drag force; compressible	5	0.729	0.137	0.052	0.228	0.156	5	1
Q20	Boundary layer; compressible	6	0.729	0.332	0.179	0.266	0.169	5	2
Q21	Boundary layer; incompressible	6	0.733	0.682	0.155	0.206	0.107	4	1
Q22	Continuity; incompressible	4	0.723	0.507	0.452	0.371	0.273	4	0
Q23	Continuity/Bernoulli; incompressible	3	0.720	0.597	0.456	0.399	0.305	5	1
Q24	Boundary layer; compressible	6	0.727	0.559	0.313	0.306	0.205	5	0
Q25	Impulse-momentum; incompressible	4	0.737	0.573	0.238	0.158	0.052	5	2
Q26	Boundary layer; compressible	6	0.726	0.469	0.374	0.318	0.217	4	0
Q27	Continuity/Bernoulli; incompressible	3	0.719	0.664	0.444	0.420	0.331	5	1
Q28	Drag force; compressible	5	0.734	0.730	0.229	0.176	0.081	5	2
Q29	Drag force; compressible	5	0.723	0.569	0.409	0.363	0.266	5	2
Q30	Pressure measurement; compressible	2	0.720	0.137	0.213	0.423	0.359	5	0
Q31	Continuity/Temperature variations;	4	0.718	0.332	0.455	0.435	0.348	5	0
Q32	Fluid properties (viscosity)	1	0.730	0.896	0.125	0.184	0.119	5	4
Average			**0.724**	**0.524**	**0.345**	**0.336**	**0.247**	**4.667**	**1.067**
Range			**0.712 to 0.744**	**0.137 to 0.915**	**0.034 to 0.562**	**0.047 to 0.528**	**-0.058 to 0.456**	**4 to 5**	**0 to 4**

4) Versatile Criterion is based on the Difficulty Index and PBCC. It provides a range of recommendations from Remove to Review through Keep, favoring positive PBCC values near or greater than 0.3 and higher difficulty values. The criteria for this recommendation are based on [44] and are reproduced below. This approach helps include simple questions to motivate students at the start of a test or confirm their understanding of basic concepts. It can also retain challenging questions if they effectively distinguish between distinct levels of student performance. This Versatile Criterion may be used for multiple-choice tests such as short quizzes, midterms, and final examinations.

Table 4. Versatile Criterion showing which questions to keep, remove, and review

Difficulty Score (%)	PBCC $[0.3,\ 1.0]$	PBCC $[0.15,\ 0.3)$	PBCC $[0.0,\ 0.15)$	PBCC $[-1,\ 0)$
$[0,\ 30]$	Review	Review/Remove	Remove	Remove
$(30,\ 50]$	Keep (Tough)	Review	Review/Remove	Remove
$(50,\ 80]$	Keep	Keep	Review/Keep	Review
$(80,\ 100]$	Keep	Keep	Keep (Easy)	Review

Based on these four criteria, the recommendations for keeping or removing the FMCI questions are as follows. The Versatile Criterion option provides more granular recommendations, such as the question may be kept but is tough (low difficulty index).

Table 5. Four criteria results to show keeping or removing a question

Question	Title	Concept	Check Alpha	Check Jorion	Check Versatile	Check Stringent
Q3	Continuity; compressible	4	Keep	Keep	Keep	Keep
Q4	Bernoulli; incompressible	1	Keep	Keep	Keep	Keep
Q5	Boundary conditions	3	Keep	Keep	Keep (Tough)	Keep
Q6	Momentum; incompressible	4	Keep	Keep	Keep (Tough)	Keep
Q7	Pressure definition	1	Keep	Remove	Keep	Remove
Q8	Boundary layers; incompressible	6	Keep	Keep	Keep	Keep
Q9	Pascal's Law	2	Keep	Keep	Keep	Keep
Q10	Manometry; compressible	2	Keep	Keep	Keep (Tough)	Keep
Q11	Bernoulli; incompressible	3	Keep	Keep	Keep (Tough)	Keep
Q12	Forces on submerged surface	2	Keep	Keep	Keep	Keep
Q13	Ideal Gas Law	4	Keep	Keep	Keep (Tough)	Keep
Q14	Manometry; compressible	2	Keep	Keep	Keep (Tough)	Keep
Q15	Shear stress; compressible	1	Keep	Keep	Keep (Tough)	Keep
Q16	Boundary layers; incompressible	1	Remove	Remove	Review/Keep	Remove
Q17	Bernoulli; incompressible	3	Keep	Keep	Keep	Keep
Q18	Manometry; compressible	2	Keep	Keep	Keep (Tough)	Keep
Q19	Drag force; compressible	5	Keep	Remove	Remove/Review	Remove
Q20	Boundary layer; compressible	6	Keep	Remove	Review	Remove
Q21	Boundary layer; incompressible	6	Remove	Remove	Keep	Remove
Q22	Continuity; incompressible	4	Keep	Keep	Keep	Keep
Q23	Continuity/Bernoulli; incompressible	3	Keep	Keep	Keep	Keep
Q24	Boundary layer; compressible	6	Keep	Keep	Keep	Keep
Q25	Impulse-momentum; incompressible	4	Remove	Keep	Keep	Remove
Q26	Boundary layer; compressible	6	Keep	Keep	Keep (Tough)	Keep
Q27	Continuity/Bernoulli; incompressible	3	Keep	Keep	Keep	Keep
Q28	Drag force; compressible	5	Remove	Keep	Keep	Remove
Q29	Drag force; compressible	5	Keep	Keep	Keep	Keep
Q30	Pressure measurement; compressible	2	Keep	Remove	Review	Remove
Q31	Continuity/Temperature variations;	4	Keep	Keep	Keep (Tough)	Keep
Q32	Fluid properties (viscosity)	1	Keep	Remove	Keep	Remove
Number of keeps			**26**	**23**	**16**	**21**

Using only the Jorion et al. Criterion [5], seven of the 30 questions should be removed. These are Q7, Q16, Q19, Q20, Q21, Q30, and Q32, and they are shown with the two critical statistical measures of the classical test theory in Table 6.

Table 6. Candidates for question removal based on Jorion et al. [5] criterion

Question	Difficulty	Discrimination
Q7	0.915	0.125
Q16	0.592	0.034
Q19	0.137	0.052
Q20	0.332	0.179
Q21	0.682	0.155
Q30	0.137	0.213
Q32	0.896	0.125

As one can note, Q7 and Q32 are too easy, and Q19 and Q30 are too difficult. Q16 does not discriminate between low and high performers. The other two questions, Q20 and Q21, are at the fringes of the criterion. What are the content reasons for such removals? As per the last author of this paper, the reasons are given below without giving away the question and its correct answer.

I. Q7 (too easy): This question is considered too easy because the concept of pressure and its directional properties are thoroughly addressed in numerous prerequisite courses such as physics and thermodynamics. Additionally, this knowledge is frequently applied in almost every question on midterm and final exams, allowing students ample opportunity to practice.
II. Q16 (does not discriminate): The question needs more explicit wording since two answers might be correct. Also, the statement is lengthy and repetitive.
III. Q19 (too difficult): The question is challenging because the answer aligns differently from one's intuition.
IV. Q20 and Q21 (nearly outside the acceptable range): Q20 pertains to a compressible boundary layer, while Q21 concerns an incompressible boundary layer. Relying on intuition grounded in solid mechanics principles for compressible flow would mislead students. Additionally, the instructor does not delve deeply into boundary layer concepts, possibly because these topics are covered more comprehensively in a fourth-year Heat Transfer course.
V. Q30 (too difficult): This question is manageable, but the illustration is unnecessarily complex and includes irrelevant details, which confuses the students.
VI. Q32 (simple): This question is clear-cut, as the correct answer matches the textbook definition of viscosity. However, it can be misleading because another option might be correct based on the value of one of the dimensionless numbers used in fluid mechanics. Yet, another choice could apply to specific conditions.

We now review item response theory before removing any questions for additional analysis or providing Fluid Mechanics instructors reasons to update the FMCI.

4.5 Item Response Theory

In classical test theory, all items on a scale intended to measure a single construct are generally considered interchangeable. However, Item Response Theory (IRT) [45] evaluates the distinct characteristics of each item. Examining the resulting Item Characteristic Curves (ICCs) is a straightforward method for understanding this. These curves illustrate the likelihood of an examinee answering correctly based on their ability. This probability is small for individuals with lower ability and higher for those with greater ability. Each item produces a smooth S-shaped curve. In Figure 1, with ability ranging from -4 to +4, the probability starts near zero at lower levels and approaches one at higher levels. An S-shaped curve is considered to be an indicator of a good question. For the FMCI, as shown in Figure 1, we see Q7, Q19, Q30, and Q32 clearly showing non-S-shapes. CTT picked these questions as well.

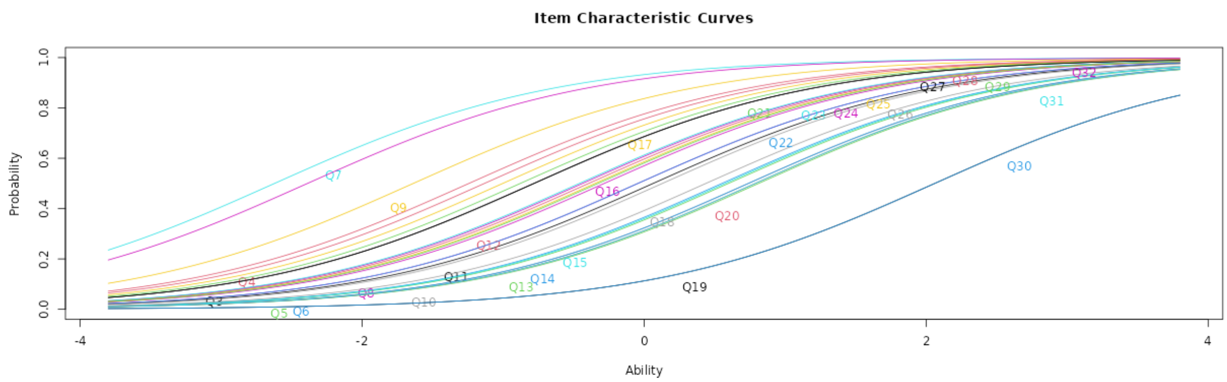

Figure 1. ICC curves for FMCI questions

Table 7 shows the subscale alpha for each of the six concept groups. The values are reasonable for concepts 2, 3, and 4, while low for 1, 5, and 6. We have removed the Q7, Q19, Q30, and Q32 from the CI statistics as these did not meet CTT and IRT expectations. The same table also shows the subscale alphas without these questions. The results for the subscale alphas are varied, with two decreasing and one increasing by a nonsignificant amount. This variation suggests that a revised FMCI may be necessary.

Table 7. Subscale alphas with and without removing questions

Concept Group	Subscale alpha with all questions	Subscale Alpha with questions removed
1	0.136	0.016
2	0.516	0.464
3	0.716	0.716
4	0.419	0.419
5	0.112	0.136
6	0.224	0.224

4.6 Correlation of CI Score to Final Exam

Figure 2 demonstrates a modest correlation of r=0.24 between the final exam percentage score and the FMCI percentage score. This indicates that merely 6% of the variability in the final exam scores can be attributed to the variation in the concept inventory. Administered at the semester's conclusion, the final exam comprehensively covered the entire course content. The final exam is a 2-hour assessment given in the last week of the semester during final exam week. The exam comprises 12 multiple-choice questions reflecting lower Bloom's taxonomy levels and three free-response questions targeting higher Bloom's taxonomy levels [46, 47]. The final examination included both types of questions because it provided a comprehensive assessment of students' understanding of the material. A recommended time distribution was 1 hour each for the multiple-choice and free-response sections, facilitating a balanced approach to test knowledge recall and analytical reasoning. The exam was conducted in a closed-book and closed-notes format. However, students were permitted a one-page, single-sided formula sheet sized 8.5"×11" to assist them while answering the questions. In contrast to the FMCI, students were allowed to use a scientific calculator. Throughout all three semesters, students were administered the same final exam to guarantee a consistent and equitable evaluation of their knowledge base. Each multiple-choice question was deliberately designed to correspond with the key concepts previously identified in the CI test (Table 1). The free-response portion required students to articulate their problem-solving process through written explanations, focusing on fluid mechanics problems.

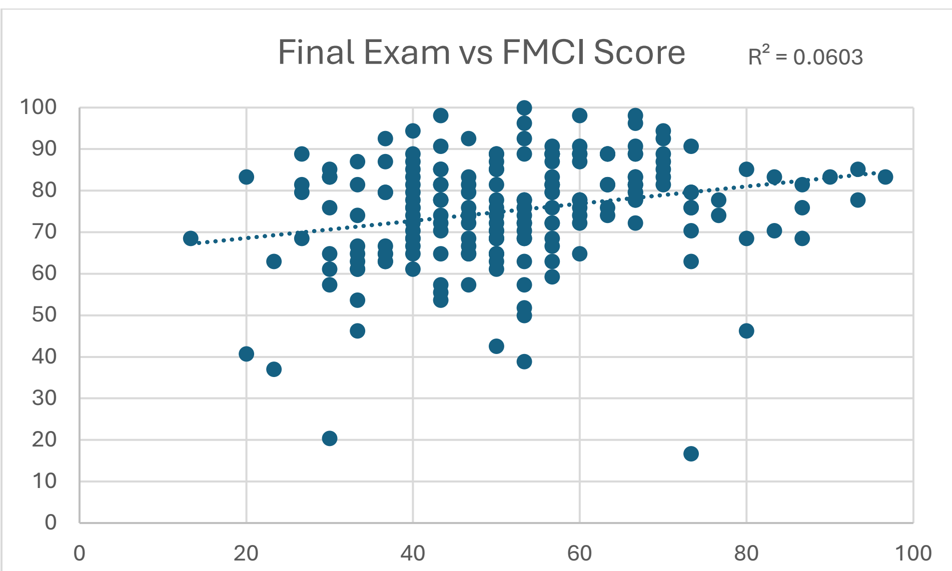

Figure 2. Final exam percentage score vs. concept inventory percentage score

5. Demographic Differences in Student Performance in FMCI

We looked at the student performance of various demographics in the FMCI. The averages of the different groups showed negligible differences, as measured by using effect size formulas [48], as shown in the table below. We could not use statistical methods such as Analysis of Means because the demographic subgroups overlap; hence, the subgroups are not independent.

Table 8. Student performance in concept inventory by demographics

Category	Subcategory	Count	Average percent	Standard Deviation	Size Difference
Overall	Overall	208	52.29	15.33	
Socioeconomic Status	Pell Grant Recipients	56	52.74	16.37	0.00
	Pell Grant Nonrecipients	141	51.58	14.58	0.00
Prerequisite GPA	Prerequisite GPA>3.5	100	53.30	12.71	0.00
	Prerequisite GPA<=3.5	97	50.48	17.13	-0.01
Gender	Male	158	51.96	14.53	0.00
	Female	34	50.59	18.10	-0.01
Age	Age >21 yrs	77	52.29	17.02	0.00
	Age<=21 yrs	118	51.50	13.73	0.00
Academic Level	4th Year Students	51	52.22	18.27	0.00
	3rd Year Students or lower	145	51.86	13.89	0.00
Transfer Status	Entered USF as a first-year Student	130	51.90	14.85	0.00
	Transferred from CC	46	51.45	15.69	0.00
	Other transfers	21	53.01	15.77	0.00
Minority Status	Underrepresented Minority	46	51.81	15.68	0.00
	Non-underrepresented Minority	122	51.94	15.03	0.00

6. Proposed Method for the Development of the Revised FMCI

Drawing from our experience proctoring FMCI to approximately 250 students and carefully studying the FMCI, we present the following recommendations. Fluid Mechanics courses play a crucial role in the curricula of Mechanical, Civil, and Environmental Engineering majors. However, there are significant differences in the focus and time allocation dedicated to various fluid mechanics concepts across these disciplines. For instance, the version of the course offered for mechanical engineering students primarily concentrates on fluid dynamics, delving deeply into the behavior of fluids in motion and their applications in engineering systems. In contrast, the Fluid Mechanics course in civil engineering prioritizes fluid statics, emphasizing the properties and behaviors of stationary fluids and their implications for structures and infrastructure. Given these differing emphases, it would be beneficial to develop specialized fluid mechanics concept inventories tailored to the unique needs of each engineering major. This targeted approach could significantly enhance the effectiveness of the CI, ensuring that students grasp the relevant fluid mechanics principles that are most applicable to their field of study.

The current Fluid Mechanics Concept Inventory (FMCI) consists of over 30 questions and is designed to be administered for 60 or 75 minutes during a lecture period. However, certain fluid mechanics courses are structured for 50 minutes, meeting thrice weekly. This discrepancy presents an opportunity to create two distinct versions of a new fluid mechanics concept inventory. The first version could assess students on six key concepts, with three closely related questions dedicated to each concept as followed in the Numerical Methods concept inventory [29]. This version would fit within a 50-minute class session. Conversely, the second version will encompass nine concepts, maintain three questions per concept, and be intended for a 75-minute class period. This time allocation would allow for a more comprehensive evaluation of the student's understanding of fluid mechanics, accommodating additional concepts within the extended timeframe.

The authors recommend using the Delphi technique [11] to develop a new FMCI. The Delphi technique is structured to gather expert opinions and achieve consensus on specific topics. It is beneficial for developing concept inventories in educational settings, where identifying key concepts and common misconceptions is crucial. The steps for following the Delphi technique [11] in creating a concept inventory are explained in detail in the paper on developing the Numerical Methods concept inventory [29].

Following the Delphi technique ensures a systematic and collaborative approach to developing a concept inventory. By leveraging the expertise of subject matter experts and achieving consensus, a reliable and valid assessment tool can be created that accurately measures students' conceptual understanding and identifies misconceptions.

7. Conclusions

This study evaluated the Fluid Mechanics Concept Inventory (FMCI), which was administered to students in a mandatory Fluid Mechanics course at a large southeastern university. The FMCI, which consists of 30 multiple-choice questions, was used to assess students' understanding of key fluid mechanics concepts. Our analysis revealed several insights into the reliability and quality of the inventory.

The calculated Cronbach's alpha of 0.73 indicates a reasonable level of reliability for the FMCI. However, the weak correlation between the concept inventory scores and final exam scores suggests that the FMCI may not fully capture students' comprehensive understanding of the course material. Several questions were identified as candidates for removal based on various criteria, including item difficulty, item discrimination, and item response theory. Specifically, questions that were either too easy, too difficult, or did not effectively discriminate between high and low performers were recommended for removal.

The limitations of the current FMCI highlight the need for a revised inventory that better aligns with the specific concepts taught in the course. The Delphi method is suggested as a systematic approach to developing a more effective FMCI. This revised inventory should ensure that each question accurately measures students' understanding of the targeted concepts, enhancing the overall assessment process.

In conclusion, while the current FMCI provides valuable insights into students' conceptual understanding of fluid mechanics, there is room for improvement. A revised and more targeted concept inventory will not only improve the reliability and validity of the assessment but also contribute to better teaching strategies and enhanced student learning outcomes in fluid mechanics.

Acknowledgments

This material is based on work partially supported by the National Science Foundation under Grant Numbers 2019664 and 2335800. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Dr. Renee Clark of the University of Pittsburgh for providing the demographic data collected during the study and Dr. J.K. Martin for providing the FMCI. Readers can obtain the concept inventory by emailing him at martin@engr.wisc.edu.

References

1. I. W. Wait and E. J. Nelson, "Concept inventory for engineering hydrology–development and implementation," in *ASEE Annual Conference & Exposition,* Seattle, WA, 2015, pp. 26.389. 1-26.389. 20.

2. J. Richardson, P. Steif, J. Morgan, and J. Dantzler, "Development of a concept inventory for strength of materials," in *33rd ASEE/IEEE Frontiers in Education*, Boulder, CO, 2003.

3. M. Prince and M. Vigeant, "Using inquiry-based activities to promote understanding of critical engineering concepts," in *ASEE Annual Conference and Exposition,* Chicago, IL, 2006.

4. K. E. Wage, J. R. Buck, M. A. Hjalmarson, and J. K. Nelson, "Signals and systems assessment: Comparison of responses to multiple choice conceptual questions and open-ended final exam problems," in *Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, 2011, pp. 198-203, IEEE.

5. N. Jorion, B. D. Gane, K. James, L. Schroeder, L. V. DiBello, and J. W. Pellegrino, "An analytic framework for evaluating the validity of concept inventory claims," *Journal of Engineering Education,* vol. 104, no. 4, pp. 454-496, 2015.

6. B. Rittle-Johnson and R. S. Siegler, "The relation between conceptual and procedural knowledge in learning mathematics: A review," in *The Development of Mathematical Skills,* C. Donlan, Ed. Psychology Press/Taylor & Francis, 1998, pp. 75-110.

7. H. A. Schweingruber, N. R. Nielsen, and S. R. Singer, Eds. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. National Academies Press, 2012.

8. D. Montfort, "An investigation of students' conceptual understanding in related sophomore to graduate-level engineering and mechanics courses," M.S. thesis, Department of Civil and Environmental Engineering, Washington State University, 2007.

9. J. K. Martin, J. Mitchell, and T. Newell, "Work in progress: Analysis of reliability of the fluid mechanics concept inventory," in *34th ASEE/IEEE Frontiers in Education*, Savannah, GA, 2004.

10. J. Martin, J. Mitchell, and T. Newell, "Development of a concept inventory for fluid mechanics," in *33rd ASEE/IEEE Frontiers in Education,* Boulder*,* CO*,* 2003.

11. R. A. Green, "The Delphi technique in educational research," *Sage Open,* vol. 4, no. 2, Article 2158244014529773, 2014.

12. R. A. Streveler, B. M. Olds, R. L. Miller, and M. A. Nelson, "Using a Delphi study to identify the most difficult concepts for students to master in thermal and transport science," in *Proceedings of the Annual Conference of the American Society for Engineering Education,* Nashville, TN, 2003.

13. G. Y. Grabarnik, M. Guysinsky, and S. Yaskolko, "An integrated approach to using statistics concept inventories in instruction," in *4th IEEE Integrated STEM Education Conference*, Princeton, NJ, 2014.

14. C. A. R. García, J. L. B. Pérez, and J. L. R. Ochoa, "Correlation between procedural and conceptual test in a statics course," in *IEEE Frontiers in Education Conference (FIE)*, San Jose, CA, 2018.

15. G. Rowe and C. Smaill, "Development of an electromagnetics course concept inventory," in *Annual Conference and Exposition of the American Society for Engineering Education,* Pittsburgh, PA, 2008, pp. 13.418. 1-13.418. 13.

16. J. Richardson, "Concept inventories: Tools for uncovering STEM students' misconceptions," *Assessment and Education Research,* pp. 19-26, 2004.

17. D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *The Physics Teacher,* vol. 30, no. 3, pp. 141-158, 1992.

18. I. A. Halloun and D. Hestenes, "The initial knowledge state of college physics students," *American Journal of Physics,* vol. 53, no. 11, pp. 1043-1055, 1985.

19. S. Krause and A. Tasooji, "Diagnosing students' misconceptions on solubility and saturation for understanding of phase diagrams," in *Annual Conference and Exposition of the American Society for Engineering Education,* Honolulu, HI, 2007.

20. K. Allen, "The statistics concept inventory: Development and analysis of a cognitive assessment instrument in statistics" Ph.D. dissertation, University of Oklahoma, Norman, OK, 2006.

21. D. Sands, M. Parker, H. Hedgeland, S. Jordan, and R. Galloway, "Using concept inventories to measure understanding," *Higher Education Pedagogies,* vol. 3, no. 1, pp. 173-182, 2018.

22. S. Y. Nozaki, S. A. Sorby, N. E. Study, H. M. Steinhauer, M. Sadowski, and R. Miller, "The development of a concept inventory for engineering graphics," in *IEEE Frontiers in Education Conference (FIE)*, Erie, PA, 2016.

23. "Concept Inventory Assessment Tools." Accessed: April 8, 2025. Available: https://www.foundationcoalition.org/home/keycomponents/concept/index.html#1

24. R. Miller, R. Streveler, M. Nelson, M. Geist, and B. Olds, "Concept inventories meet cognitive psychology: Using beta testing as a mechanism for identifying engineering student misconceptions," in *Annual Conference and Exposition of the American Society for Engineering Education,* Portland OR, 2005, pp. 10.335. 1-10.335. 18.

25. B. M. Olds, R. A. Streveler, R. L. Miller, and M. A. Nelson, "Preliminary results from the development of a concept inventory in thermal and transport science," in *Annual Conference and Exposition of the American Society for Engineering Education,* Salt Lake City, UT, 2004.

26. K. Allen, A. Stone, T. R. Rhoads, and T. Murphy, "The statistics concepts inventory: Developing a valid and reliable instrument," in *Annual Conference and Exposition of the American Society for Engineering Education*, Salt Lake City, UT, 2004.

27. A. Stone, K. Allen, T. R. Rhoads, T. J. Murphy, R. L. Shehab, and C. Saha, "The statistics concept inventory: A pilot study," in *33rd ASEE/IEEE Frontiers in Education,* Boulder, CO, 2003.

28. P. S. Steif and J. A. Dantzler, "A statics concept inventory: Development and psychometric analysis," *Journal of Engineering Education,* vol. 94, no. 4, pp. 363-371, 2005.

29. A. K. Kaw, Y. Lou, A. Scott, and R. L. Miller, "Building a concept inventory for numerical methods: A chronology," in *Annual Conference and Exposition of the American Society for Engineering Education,* New Orleans, LA, 2016.

30. D. R. Webster, D. M. Majerich, and A. G. Madden, "Flippin' fluid mechanics - comparison using two groups," *Advances in Engineering Education,* vol. 5, no. 3, 2016.

31. D. Evans *et al.*, "Progress on concept inventory assessment tools," in *33rd ASEE/IEEE Frontiers in Education,* Boulder, CO, 2003.

32. M. K. Watson, A. R. Mills, K. C. Bower, K. Brannan, M. H. Woo, and R. W. Welch, "Refinement of a concept inventory to assess conceptual understanding in civil engineering fluid mechanics," in *Annual Conference and Exposition of the American Society for Engineering Education,* Seattle, WA, 2015.

33. D. G. E. Huilier, "Forty years' experience in teaching fluid mechanics at Strasbourg University," *Fluids,* vol. 4, no. 4, 2019.

34. D. Fraser, R. Pillay, L. Tjatindi, and J. Case, "Enhancing the learning of fluid mechanics using computer simulations," *Journal of Engineering Education,* vol. 96, no. 4, pp. 381-388, 2007.

35. A. W. Baghdanov, "Student understanding of pressurized pipeline concepts and the application to conceptual change theory," M.S. thesis, Department of Civil & Environmental Engineering, Washington State University, 2013.

36. S. A. Brown and M. S. Barner, "Board# 13: Examining engineering concepts in practice: Is conceptual understanding relevant to practice?," in *Annual Conference and Exposition of the American Society for Engineering Education,* Columbus, OH, 2017.

37. R. Guldiken, A. Kaw, and R. Clark, *Collaborative Proposal: Structured Use of Metacognitive Activities in a Flipped Undergraduate Engineering Course to Enhance Learning and Professional Skill Development*. 2020, National Science Foundation.

38. R. M. Clark, A. Kaw, and R. Guldiken, "Metacognition instruction and repeated reflection in a fluid mechanics course: Reflective themes and student outcomes," *International Journal of Mechanical Engineering Education,* vol. 51, no. 4, pp. 243-269, 2023.

39. R. M. Clark, R. Guldiken, A. Kaw, and O. Uyanik, "The case for metacognition support in a flipped STEM course," *International Journal of Mechanical Engineering Education,* Article 03064190241255113, 2024.

40. G. Aden-Buie, A. Kaw, "MC Test Analysis." Accessed: April 8, 2025. Available: https://pkg.garrickadenbuie.com/mctestanalysis/index.html, 2017.

41. M. D. Engelhart, "A comparison of several item discrimination indices," *Journal of Educational Measurement,* vol. 2, no. 1, pp. 69-76, 1965.

42. T. Haladyna and M.C. Rodriguez, *Developing and Validating Test Items*. Routledge, 2013.

43. G. Gray, F. Costanzo, D. Evans, P. Cornwell, B. Self, and J. Lane, "The dynamics concept inventory assessment test: A progress report and some results," in *Annual Conference and Exposition of the American Society for Engineering Education,* Portland, OR, 2005.

44. Sleeper, R. *Keep, Toss or Revise? Tips for Post-Exam Item Analysis.* 2011 [cited 2015 not available; Available from: http://www.ttuhsc.edu/sop/administration/enhancement/documents/Sleeper_Handout.ppt

45. S. E. Embretson and S. P. Reise, *Item Response Theory for Psychologists*. Psychology Press, 2013.

46. D. Krathwohl, "A revision of Bloom's taxonomy: An overview," *Theory into Practice,* vol. 41, no. 4, pp. 212-218, 2002.

47. B. S. Bloom, *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. Addison-Wesley Longman, 1984.

48. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 2013.

49. R. A. Streveler, R. L. Miller, A. I. Santiago-Román, M. A. Nelson, M. R. Geist, and B. M. Olds, "Rigorous methodology for concept inventory development: Using the 'assessment triangle' to develop and test the thermal and transport science concept inventory (TTCI)," *International Journal of Engineering Education,* vol. 27, no. 5, pp. 968-984, 2011.

50. K. A. Dell and G. A. Wantuch, "How-to-guide for writing multiple choice questions for the pharmacy instructor," *Currents in Pharmacy Teaching and Learning,* vol. 9, no. 1, pp. 137-144, 2017.

51. S. M. Case and D. B. Swanson, *Constructing Written Test Questions for the Basic and Clinical Sciences*. National Board of Medical Examiners, Philadelphia, 1998.

52. C. Brame. "Writing Good Multiple Choice Test Questions." Accessed: April 8, 2025. Available: https://sites.google.com/view/cynthia-brame/teaching-guides/writing-good-multiple-choice-test-questions