2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #34218

What Sticks When the Dust Settles: Evaluating the Retention of Concepts and Thought Processes with Think-aloud Interviews

Dr. Soheil Fatehiboroujeni, Cornell University

Soheil Fatehiboroujeni received his Ph.D. in Mechanical Engineering from the University of California, Merced in 2018. As a postdoctoral researcher at Cornell University, Sibley School of Mechanical and Aerospace Engineering, Soheil is working in the Active Learning Initiative to promote student learning and the use of computational tools such as Matlab and ANSYS in the context of fluid mechanics and heat transfer.

Dr. Matthew Jordan Ford, Cornell University

Matthew Ford received his bachelor's degree in mechanical engineering and materials science from the University of California, Berkeley, and went on to complete his Ph.D. in mechanical engineering at Northwestern University. After completing an internship in quantitative methods for education research with the Center for the Integration of Research, Teaching, and Learning (CIRTL), he joined the Cornell Active Learning Initiative as a postdoctoral associate. His teaching interests include solid mechanics, engineering design, and inquiry-guided learning.

Dr. Hadas Ritz, Cornell University

Hadas Ritz is a senior lecturer in Mechanical and Aerospace Engineering, and a Faculty Teaching Fellow at the James McCormick Family Teaching Excellence Institute (MTEI) at Cornell University, where she received her PhD in Mechanical Engineering in 2008. Since then she has taught required and elective courses covering a wide range of topics in the undergraduate Mechanical Engineering curriculum. In her work with MTEI she co-leads teaching workshops for new faculty and assists with other teaching excellence initiatives. Her main teaching interests include solid mechanics and engineering mathematics. Among other teaching awards, she received the 2020 ASEE St. Lawrence Section Outstanding Teaching Award.

Prof. Elizabeth Mills Fisher, Cornell University

Elizabeth M. Fisher is an Associate Professor in the Sibley School of Mechanical and Aerospace Engineering at Cornell. She received her PhD from U.C. Berkeley.

WIP: What sticks when the dust settles: evaluating the retention of concepts and thought processes with think-aloud interviews

A foundational goal of deliberately designed educational experiences is for learners to acquire and to retain knowledge and thinking skills. Common modes of assessment such as exams administered during the semester may measure students' short-term achievement of learning objectives. However, they do not identify the underlying knowledge gaps, misconceptions, or thought processes responsible for students' errors. In addition, such tests do not yield information on students' retention of skills beyond the conclusion of the course. Such limitations make it difficult to optimize learning activities for long-term retention of concepts and skills that are crucial to problem solving activities in engineering courses. Think-aloud interviews provide rich qualitative data about students' thought processes in verbal form that are not available from multiple-choice or free-response type of assessments. In this work-in-progress paper we present a set of think-aloud interviews with senior-level mechanical engineering students as they take a concept inventory assessment on introductory fluid mechanics and mechanics of engineering materials. Participants took these two courses in their junior year, at the end of which they took the same concept inventory tests, but not in a think-aloud format. We explain how our interview data is used to identify students' knowledge gaps or misconceptions.

1. Introduction

From a pedagogical point of view, learning experiences are designed to foster specific skills and dispositions that students demonstrably achieve in the course of their learning. For students' success, it is also critical that skills and learning dispositions are carried with them as they progress along their degree programs and eventually into professional practice. Success in teaching and learning, in this regard means *make it stick* [1]. Due to practical limitations it is not common to directly assess students' retention of knowledge when significant time has passed (for instance one academic year) from the time they take courses. Yet such assessments of retention could provide valuable insights for the design of learning activities that support *lasting* student achievements.

Long term retention of knowledge and skills are known to be impacted by a range of factors which themselves may interact [2, 3]. The main factors that are most often reported in the cognitive and educational research [2] are as follows: (1) the degree of original learning, (2) task characteristics such as complexity, (3) the size of the retention interval, (4) pedagogy and instructional strategies, (5) testing methods, (6) individual differences. Some early studies have shown [3] that the retention of conceptual knowledge from a highly structured knowledge domain, such as theories or models in natural sciences, should show less rapid losses than more detailed and less schema-central domains such as the names of different species and chemical compounds. While researchers of mathematics education [4] have studied how instructional approaches can impact the long term retention of both conceptual and procedural learning, in engineering education there have not been many published reports available on the cognitive retention of students after the passage of significant time, e.g. several months. This paper is a work-in-progress report on a small step we are taking in this direction within the Cornell Active Learning Initiative (ALI) at the department of mechanical and aerospace engineering.

Recently we developed a set of multiple choice learning assessments for a number of courses in the mechanical engineering major [5] which includes introduction to fluid mechanics and mechanics of engineering materials. These assessments were administered at the end of Fall 2019 to students in each course. Students were informed of the purpose of these assessments and were incentivized to participate in them by a completion-based grade (i.e. not based on the correctness of their responses). About one year later, in Fall 2020 we used think-aloud interviews [6] to observe the performance of students as they completed the same learning assessments. In the following sections we first put into perspective our evolving research question and then report on some of the preliminary results of our data analysis.

Research Question: what are the course concepts that students are less likely to recall or to successfully apply in problem solving activities one year after taking a course?

The thinking required in problem solving activities is often mapped into cognitive tasks with varying degrees of complexity. Recalling or comprehending facts, ideas and concepts are cognitively simpler than deriving implications from the facts or applying the concepts and ideas in solving a problem. Moreover, students in accordance with their prior experiences, level of engagement, and achievement at the time of taking a course, may be familiar with the thinking involved in the solving of certain problems to varying degrees. When faced with problems that they do not just recall how to solve from their past experiences, students need a type of knowledge transfer in order to construct their solution without simply recalling the necessary solution steps. As a point of departure, we use Bloom's taxonomy [7, 8] as a framework to identify the basic types of abilities and skills that students exhibit in problem solving activities during the think-aloud interviews.

2. Methods

2.1. Data collection and interview protocol

In Fall 2020, out of the students who took both courses (introduction to fluid mechanics and mechanics of engineering materials) we invited those who had submitted the multiple choice learning assessments in both courses to participate in a think-aloud study. We incentivized their participation with a small financial reward. We combined selected questions from the two learning assessments so as to (a) develop a combined assessment that can be finished in a think-aloud interview within an hour and (b) have all the key topical areas in each course covered in the new subset. The original fluid mechanics and mechanics of material assessments in 2019 had 36 and 20 problems, respectively. In 2020 we combined 11 of the fluids with 19 of the solids problems into one assessment for the think-aloud studies.

Two researchers in our team each conducted four interviews for a total of eight. To define a measure of how our participants represent the population we divided the students into four quadrants based on their scores in the 2019 assessments and selected two students from each quadrant. We divided the potential participants into four groups based on their performance on the two learning assessments, i.e. high fluids score, high solids score; high fluids score, low solids score, etc. The cutoffs for high- and low-scorers were slightly adjusted from the medians to roughly balance the number of potential participants in each quadrant. We produce a transcript of the interviews that are coded as we explain in section 2.3, but we first give a description of how we operationalize the categories of Bloom's taxonomy.

2.2. Mapping our assessment items onto Bloom's taxonomy:

Here we present the descriptions of the cognitive activities and skills that we adopt from Bloom's taxonomy. It is important to note that Bloom's categories follow a specific type of hierarchy. The simpler tasks such as recalling are subsumed as a necessary but not sufficient component of the more complex cognitions such as application or analysis. So an important characteristic of the Bloom's system is its qualitative distinction between certain *levels* of activity that are built on top of one another in a way that lower levels are both surpassed and conserved by the higher levels.

I. Knowledge recall: This activity refers to "the recall of specifics and universals, the recall of methods and processes, or the recall of a pattern, structure, or setting." [7] Although the "processes of remembering*"* is mainly emphasized in an activity which requires knowledge retention, learners also need to establish certain relationships to recall relevant knowledge during the problem solving. As it is described by Bloom [7] "the process of relating is also involved in that a knowledge test situation requires the organization and reorganization of a problem such that it will furnish the appropriate signal." We do not include problems that exclusively test this skill, however we frequently observe this as a type of thinking activity that students use to answer questions.

II. Comprehension: This refers to the most basic level of *understanding* such as translating, interpreting, or extrapolating given information whether it is verbal, graphical, or mathematical. When an "individual knows what is being communicated and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications" [7] we have sufficient evidence for their comprehension skills being demonstrated.

Example 1. From the mechanics of materials assessment: When a metal wire is pulled through a hole smaller than its initial diameter, its strength increases. This is primarily because:

A. the material has fewer dislocations.

B. the material has more dislocations.

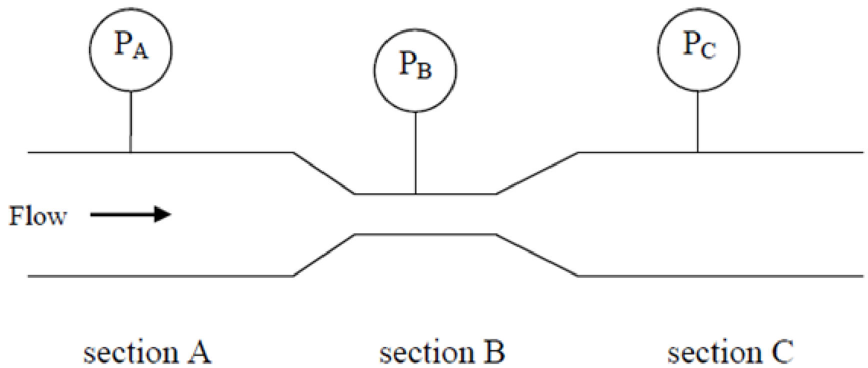
C. the compaction of the atoms increases the strength of the interatomic bonds.

D. the compaction of the crystals increases the strength of the grain boundaries.

E. the wire has been heated by friction through the die.

III. Application: The *use* of abstractions such as equations, relationships, or generalizable concepts in particular and concrete situations defines this activity. Bloom gives the following detail with regard to the demarcation between comprehension and application. "Given a problem new to the student, he [sic] will apply the appropriate abstraction without having to be prompted as to which abstraction is correct or without having to be shown how to use it in that situation. A demonstration of 'Comprehension' shows that the student *can* use the abstraction when its use is specified. A demonstration of 'Application' shows that he *will* use it correctly, given an appropriate situation in which no mode of solution is specified" [7] - italic emphasis here is added. The following problem is an example from our learning assessment that gives students the opportunity to demonstrate this level of thinking.

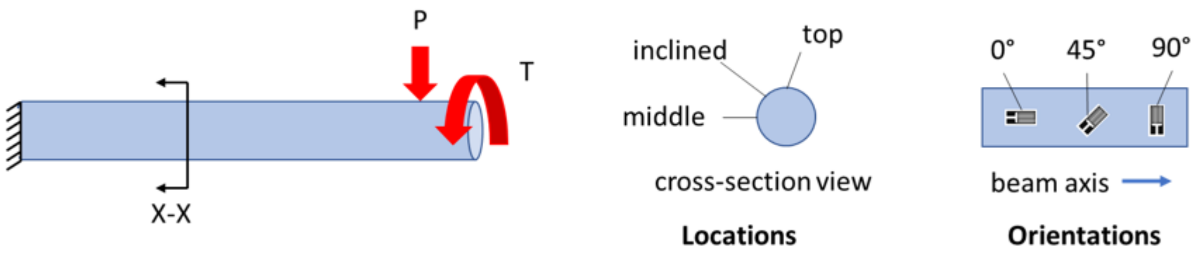
Example 2. From the fluid mechanics assessment: For the piping system shown below, water is flowing from left to right at steady state with uniform temperature. The pipe diameter is larger in section A than section B. The diameters of sections A and C are the same. If gravity and viscous effects are negligible, which of the following relationships is true about the static pressure in sections A and B? (Adopted from the literature [9].)



A. $P_B < P_A$ B. $P_B = P_A$ C. $P_B > P_A$ D. There is not enough information provided.

IV. Analysis: This refers to the breakdown of a problem into its constituent parts "such that the relative hierarchy of ideas is made clear and/or the relations between the ideas expressed are made explicit." [7] In application, students must recognize which is the relevant abstraction that needs to be put to use to solve the problem and thus what is at stake is a many-to-one type of mapping between the given information and the underlying abstraction that ought to be used. In analysis, on the other hand students require some ground or supposition to first determine the set of features or aspects of the problem that must be taken into account so as to solve the problem correctly and as such a many-to-many mapping here is involved. When students are faced with problems that require a judicious application of several interconnected rules or concepts (e.g. the derivation of stresses at multiple points on a member subject to combined loading, as well as the application of stress transformations for each element which is required in the following example) then their analytic skills can be tested.

Example 3. From the mechanics of material assessment: A circular bar will be tested simultaneously in bending and torsion, as shown below: A strain gauge measures strain along its length. Two strain gauges will be placed somewhere at cross-section X-X. Where should the two strain gauges be placed such that the bending strain and torsion strain can easily be measured independently?



A. Middle 90° and inclined 0°

B. Top 90° and inclined 45°

C. Top 0° and middle 90°

D. Middle 0° and inclined 45°

E. Top 0° and middle 45°

2.3. Coding Process

Our first step in data analysis was exploratory coding that aimed to find stable assignment of the foregoing codes to our data. Two coders went through two interview transcripts and individually coded them. Next, they reviewed and discussed any discrepancies and disagreements together about the operationalization of Bloom's categories to arrive at resolutions. During these steps each coder had also developed their own set of notes and observations in the form of emergent categories informed by thematic analysis [10] that went beyond that of Bloom's categories. These emergent codes were also discussed to achieve stable definitions. After this step, one of the coders continued with another two transcripts to cover a total of four interviews. To keep the following section brief we only present the results of our preliminary themes that are derived from the structure shown in Figure 1.

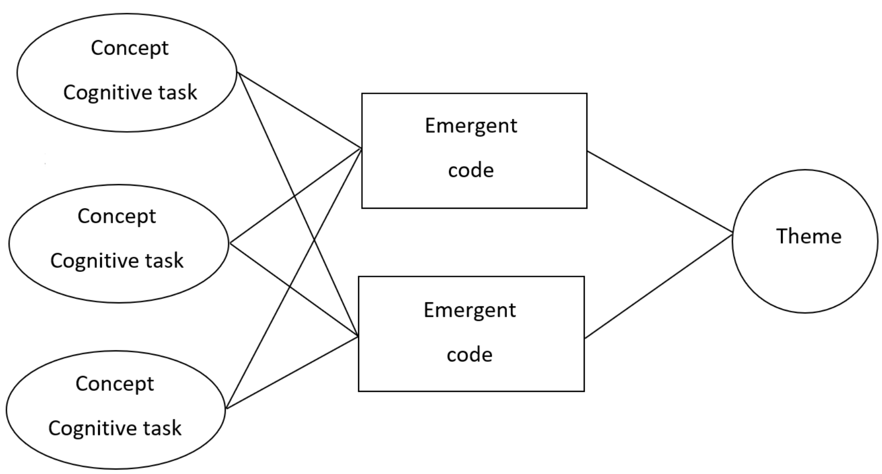


Figure 1. The figure shows a schematic concept map of our data analysis. *Concepts* and *cognitive tasks* are the first set of codes that respectively capture the content of students' speech and their cognition according to Bloom's taxonomy. The next layer of codes are *emergent* descriptions of the students' thinking during the problem solving. Finally *themes* are inferences we make about the students' thinking which intend to explain why they succeed or fail.

3. Results

3.1 Quantitative patterns

Figure 2 represents the changes that we observe in students' responses from 2019 to 2020 on a question-by-question basis. We can see in this figure that students' responses are moving in both directions: in some cases they correct a previously incorrect answer, and in some cases the reverse does happen. By looking at the data on a student-by-student basis, shown in Table 1, we find that total scores do not significantly change and the difference in mean score from 2019 to 2020 is 0.125 indicating a very small (near zero) gain. This could be attributed to the familiarity of students with the problems despite the passage of several months. On the other hand, Figure 2 shows that in a number of questions a significantly large net change is observable in the negative direction where a notable proportion of students change their response from correct to incorrect. On the following items more than 3 students changed their answer from correct to incorrect (red bars are larger than three units):

- FM03: Finding the direction of a pipe flow based on two measurements of the pressure.
- MM02: Comparing the strain energy of two rods under torsion.
- MM13: Using stress transformation concepts to find principal stresses on an element.

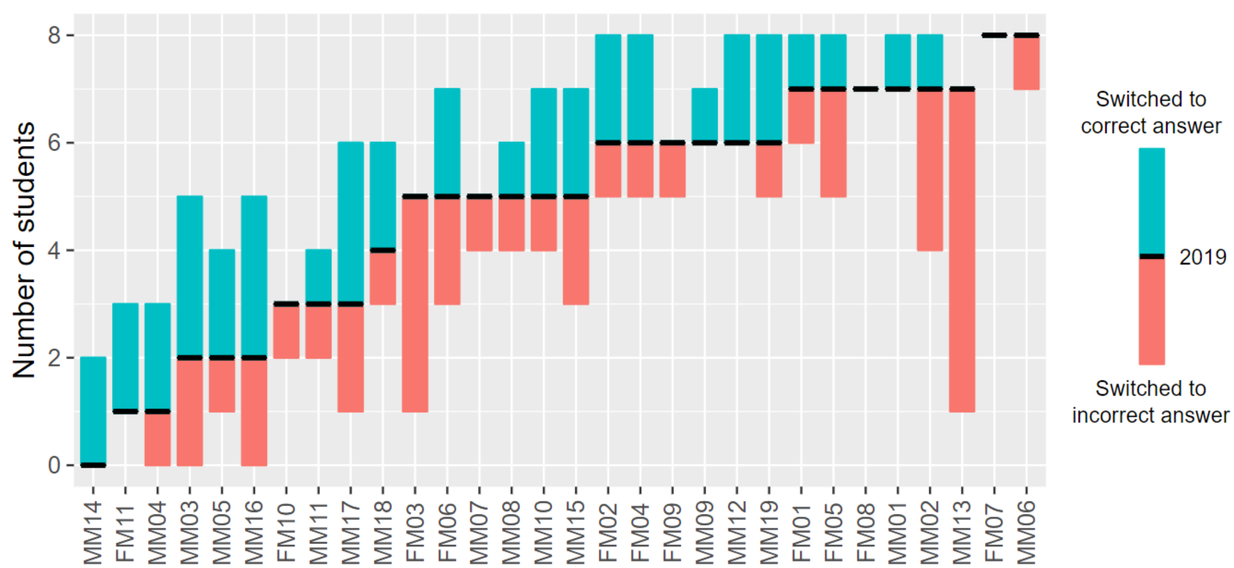


Figure 2. Changes of students' responses to each question from 2019 to 2020 are depicted by red bars if students answer incorrectly to a question that they previously got correct, and blue bars if vice versa. The black dashes are the number of students answering a question correctly in 2019 which is used to sort the data. (MM: Mechanics of Material, FM: Fluid Mechanics)

It is worth noting that there are no questions with large, obvious shifts in the positive direction. In the following topics 3 out of 8 students changed their answer from incorrect to correct while in each case, 2 students changed their answer to the incorrect. So overall these questions do not exhibit large positive shifts.

- MM03: Comparing the strain energy of two beams under bending.
- MM16: Determining the impact of the stress amplitude on the fatigue life of a shaft.
- MM17: Deciding how to increase the fatigue life span more effectively either by changing mean stress, the alternating stress, or the frequency.

To better determine the underlying reasons for these types of changes we then turn to a qualitative analysis of our data, although we have not finished this work to present any conclusive interpretation of the changes.

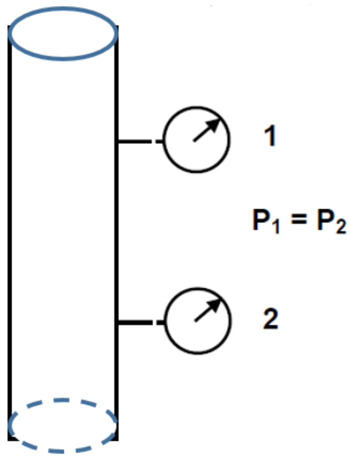
Table 1. Changes in learning assessment scores from 2019 to 2020 are captured by the following categories. CC: from correct to correct, CI: from correct to incorrect, IC: from incorrect to correct, II: from incorrect to incorrect.

Student	**quadrant**	**CC**	**CI**	**IC**	**II**	**2019**	**2020**	**net change**
S 01	A	15	5	2	8	20	17	-3
S 02	A	12	6	5	7	18	17	-1
S 03	B	12	5	7	6	17	19	2
S 04	B	13	7	6	4	20	19	-1
S 05	C	18	5	3	4	23	21	-2
S 06	C	13	5	6	6	18	19	1
S 07	D	9	5	7	9	14	16	2
S 08	D	14	1	4	11	15	18	3

3.2. Fluid Mechanics Qualitative Themes

One of our preliminary findings is that problems which involve an analysis of *combined effects of multiple phenomena* are more challenging for students. For instance, to solve the following example, students need to examine the combined effects of (a) an imposed pressure gradient that drives the flow, (b) hydrostatic changes and (c) viscous pressure drop. It is also noteworthy that the figure and the language used in the problem statement prompt students towards hydrostatic effects and pressure-induced flow, but students in all four cases that we coded so far fail to account for the pressure drop due to viscous dissipation, hence are not able to identify the correct answer to this problem. Put differently, the underlying cause of error is that viscous effects remain hidden from students' eyes in this problem.

Problem FM03: A long vertical pipe filled with water is equipped with pressure gauges at each end as shown below. An engineering student observing the pipe notices that each pressure gauge reads the same value. Is water flowing in the pipe? If so, in what direction is the water flowing? (Adopted from the literature [9].)



A. Water is flowing up.

B. Water is flowing down.

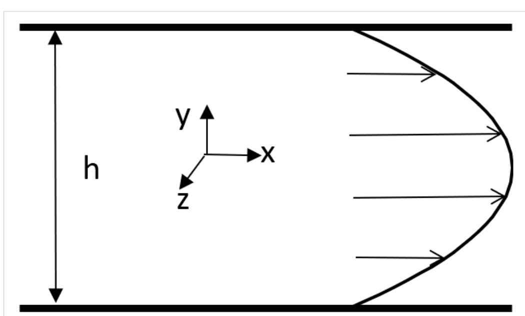
C. Water is flowing, but there is not information to decide whether it is flowing up or down.

D. Water is not flowing.

E. There is not enough information to decide whether it is flowing or not.

Another preliminary observation is that some students face a challenge in solving problems that involve the *concept of shear stress and its relationship to velocity and the fluid's viscosity*. When faced with the problem shown below, some students rely on the proportionality of the shear stresses with the velocity gradient in a Newtonian fluid in order to answer the question. Also there are cases in which students without recalling this relationship end up with guesses or confusions that lead to wrong answers.

Problem FM05: Water flows steadily through a gap of height h between two parallel plates. The velocity profile is parabolic, as shown below. At what location is the magnitude of the shear stress in the direction of flow a maximum? (Adopted from the literature [11])



A. At the centerline

B. As close as possible to one of the plates

C. In the region between the centerline and the plate

D. There is not enough information provided

3.3. Mechanics of Material Qualitative Themes

In the mechanics of material we also arrive at a similar finding with regard to *problems that draw simultaneously on multiple concepts*. Students tend to do better when tested separately on stress transformation or on the distribution of stresses in bending or torsion. But when they are asked a problem which requires the concomitant application of these two concepts such as example 3 in the previous section, students tend to face difficulties in thinking through the solution, e.g. identifying how they can independently measure torsional and bending strains in a combined loading scenario.

4. Discussion

Our ongoing coding and analysis of the think-aloud interviews so far has allowed us to identify some of the topics as well as forms of thinking that students are likely to struggle with one year after the experience of taking a course. However, there are several other steps that we are currently working on to respond to our research question. First, we intend to carry on the data analysis for emergent themes in the other half of our interview data. Next we intend to develop an inventory of all concepts included in our assessments to determine both areas in which students fail and those that they succeed. Finally we intend to further use quantitative analysis to compare the scores of 2019 with 2020. With these steps done in the near future we hope to bring further clarity to our observations and within the limits of our study design we intend to generate practical insights on the improvement of teaching and learning in these courses. We also plan to carry on this study in Fall 2021 and collect further data from students who have taken our courses and their assessments for the first time in Fall 2020.

References

[1] Brown, P. C., McDaniel, M. A., Roediger, H. L. *Make It Stick*. United Kingdom: Harvard University Press, 2014.

[2] Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General, 120*(4), 395–409. https://doi.org/10.1037/0096-3445.120.4.395

[3] Farr, M. J. (1987). *The Long-Term Retention of Knowledge and Skills: A Cognitive and Instructional Perspective (Recent Research in Psychology)* (1987th ed.). Springer. https://doi.org/10.1007/978-1-4612-1062-7

[4] Gamer, B.E., Gamer, L.E. Retention of concepts and skills in traditional and reformed applied calculus. *Math Ed Res J* 13, 165–184 (2001). https://doi.org/10.1007/BF03217107

[5] Ford, M. J., Ritz, H., Finio, B., Fisher, E. M. (2020, June), *Multiple-choice Learning Assessments for Intermediate Mechanical Engineering Courses: Insights from Think-aloud Interviews* Paper presented at 2020 ASEE Virtual Annual Conference Content Access, Virtual Online. 10.18260/1-2--34990

[6] Van Someren, M. W., Barnard, Y.F., Sandberg, J.A.C. *The think aloud method: A practical guide to modelling cognitive processes*. San Diego: Academic Press, 1994.

[7] Bloom, B. S., Masia, B. B., Krathwohl, D. R. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. United Kingdom: Longmans, Green, 1956.

[8] David R. Krathwohl "A Revision of Bloom's Taxonomy: An Overview", *Theory Into Practice*, 41:4, 212-218, 2002. doi: 10.1207/s15430421tip4104_2

[9] Ronald L. Miller, Ruth A. Streveler, Dazhi Yang, et al., "Fundamental Research in Engineering Education. Identifying and Repairing Student Misconceptions in Thermal and Transport Science: Concept Inventories and Schema Training Studies". *Chemical Engineering Education*, v. 45, n. 3, pp. 203-210, 2011.

[10] V. Braun, V. Clarke, "Thematic analysis", in *APA handbook of research methods in psychology*, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological, H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher, Eds. American Psychological Association, 2012. https://doi.org/10.1037/13620-004

[11] M. Koretsky, J. Falconer, B. Brooks, D. Gilbuena, D. Silverstein, C. Smith, and M. Miletic, "The AIChE Concept Warehouse: A Web-based Tool to Promote Concept-based Instruction", *Advances in Engineering Education*, 2014. https://eric.ed.gov/?id=EJ1076138