



**2023: VOLUME 11 ISSUE 1** DOI: 10.18260/3-1-1153-36041

# A Hands-on Guided-inquiry Materials Laboratory that Supports Student Agency

MATTHEW J. FORD School of Engineering & Technology University of Washington, Tacoma

SOHEIL FATEHIBOROUJENI Department of Mechanical Engineering Colorado State University, Fort Collins

ELIZABETH M. FISHER

HADAS RITZ Sibley School of Mechanical & Aerospace Engineering Cornell University

# ABSTRACT

Under the new ABET accreditation framework, students are expected to demonstrate "an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions." Traditional, recipe-based labs provide few opportunities for students to engage in realistic experimental design or develop a sense of agency, and research has cast doubt on their pedagogical benefit. At the same time, the COVID-19 pandemic forced institutions to move to remote learning. We developed a scaffolded series of remote lab activities for an upper-division mechanics of materials course, culminating in a collaborative guided-inquiry experiment design challenge. In the first iteration of the course, we mailed kits to students with basic supplies for the at-home experiments, while in the second students were expected to make use of readily-available household items. We analyzed 36 lab reports from the second iteration of the guided-inquiry lab and identified 25 unique approaches to the design challenge, an indication of a truly open-ended activity. Student outcomes were measured by post-lab surveys of attitudes and self-efficacy, as well as a standardized conceptual learning assessment. The fraction of students endorsing statements related to a sense of agency increased dramatically over the course of the semester: from 53% to 83% for goal-setting and from 63% to 92% for choice of methods. Self-efficacy increased significantly in the primary targeted skills (designing experiments and making predictions), but there was no significant shift in skills not explicitly targeted by the guided-inquiry lab (equitable sharing of labor, expressing opinions in a group, and interpreting graphs).



Key words: Remote laboratory [syn: Virtual laboratory], Inquiry based learning, Collaborative learning, Mechanical engineering, Experimental design, ABET

#### INTRODUCTION

The instructional laboratory experience is a hallmark of the modern engineering curriculum. Engineering students typically encounter a variety of lab experiences in different contexts, often designed with different outcomes in mind including reinforcement of lecture concepts, motivation to continue in or pursue a particular major, and development of skills in instrumentation, data analysis, teamwork, and communication (Feisel and Rosa, 2005; Holmes and Smith, 2019). Feisel and Rosa (2005) emphasized the importance of experimental design, creativity, and learning from failure as important outcomes of instructional labs in engineering. More recently, the new ABET accreditation framework requires that students demonstrate "an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions" (ABET, 2018).

In mechanical engineering, civil engineering, and materials science, some of these lab experiences involve material testing, and usually use the relevant ASTM standard as a template. A traditional "recipe-based" lab guides students through a standardized experiment using well-documented methods leading to predictable results, often for the purpose of demonstrating or confirming a concept learned in lecture. The activities that students may engage in during the activity are heavily constrained by the ASTM standard, available equipment, safety considerations, and time. Similar confirmatory lab activities in introductory physics have proven ineffective at even reinforcing lecture concepts, and the limits on student agency preclude achievement of the objectives emphasized by ABET (Holmes et al., 2017; Holmes and Wieman, 2016).

The confirmation inquiry activities just described lie at the foundation of the hierarchy of inquiry proposed by Banchi and Bell (2008). In a *structured-inquiry* activity (the next level of the hierarchy), the instructor provides the question or goal and method of investigation, while students discover the result and draw conclusions. A *guided-inquiry* activity allows students to choose or design the method of investigation for a question provided by the instructor. Finally, an *open-inquiry* leaves the task of developing the question to the students.

The purpose and value of engineering laboratories is often stated in general terms like "relating theory to practice" or bringing the "real world" into the classroom (Feisel and Rosa, 2005). But surely if making connections between the textbook and physical phenomena was the only goal, then canned lecture demonstrations would be a cheaper and more efficient alternative. What distinguishes



designing and conducting experiments from watching demonstrations is the opportunity for students to make goal-directed decisions whose outcomes affect the success of the endeavor. We propose that *student agency* is a critical component of successful student-centered pedagogy (Benedict et al., 2020). We adopt the definition of agency described by Holmes et al. (2020) as goal-directed decision-making, specifically focusing on *decision-making agency*, which is highly relevant to an engineering context, as opposed to *epistemic agency* (freedom in what questions to ask or how to create new knowledge), which is more relevant to a basic science context. Learning activities which leave room for students to make decisions may support student motivation and increase buy-in, build self-efficacy, and better reflect the conditions of professional work (Kalender et al., 2021). Students engaged in more open-ended lab activities learn more, make more expertilike decisions, and develop more sophisticated attitudes about experimental science. (Holmes et al., 2014; Wilcox and Lewandowski, 2016; Lazonder and Harmsen, 2016; Ural, 2016). Activities without a strictly prescribed method can also lead to a broader diversity of solutions (Willner-Giwerc et al., 2020), which is important for activities that involve peer review or presentations.

One approach to creating more open-ended lab experiences is to simplify or remove the instructions so that students must exercise some judgment to complete the activity (Morrison, 2014). This approach can be very simple to implement, although it is important to provide additional support and guidance for students. This approach has shown mixed results in engineering contexts and more research using specific and valid assessment instruments is needed. Ritz et al. (2018) found that students who performed a tensile test, with guidance about concepts and available equipment but no step-by-step instructions, performed better on a closely-related exam question than students who were given an explicit procedure. By contrast, Halstead (2016) found no differences in performance on a post-lab quiz in an undergraduate electronics course between students who were given an explicit recipe for building and analyzing the circuit (Maximal Scaffolding) and students who were given guidance (Reduced Scaffolding). In the electronics study, the two groups were mixed together during lab sessions, so it's possible there could have been communication between Reduced Scaffolding and Maximal Scaffolding students, while in the mechanics study entire lab sections were assigned to the same treatment group. In a fluid mechanics lab course, Johnson and Morphew (2016) had all students discuss and design their own lab procedure to meet a common goal, but students in one group were subsequently given an explicit recipe to follow after proposing their own. Students who followed their own (possibly flawed) procedure made better use of lab time and earned higher scores on lab reports graded by a standardized rubric, although it's not clear whether graders were blinded to experimental condition.

The rapid shift to remote instruction in Spring 2020 in response to the COVID-19 pandemic had a particularly powerful impact on laboratory activities. Instructors moved classes online using a



variety of methods including mailing physical kits, designing simulation-based exercises, and providing experimental data. Many group-based activities were converted into individual projects and students reported difficulties getting guidance and support (Fox et al., 2020). Some instructors responded by designing guided-inquiry activities which could be completed at home, either with widely-available household supplies or mailed kits (O'Neill, 2021; Ankeny and Tresch, 2021). Kits can be expensive and resource-intensive to assemble, and may encounter difficulties in transit.

Despite ongoing concerns about new COVID-19 variants (Jaschik, 2021), many institutions returned to in-person instruction in Fall 2021 (Hartocollis, 2021), and educators are considering what aspects of remote instruction they want to keep and adapt to in-person learning. The constraints of the pandemic led to the development of low-tech substitutes for laboratory learning. We argue that the greater flexibility and student agency afforded by at-home experiments may lead to improved learning gains over traditional recipe-based materials testing labs.

Here we report on two iterations of a low-cost, scaffolded laboratory activity sequence culminating in a guided-inquiry design activity. We analyze the cognitive tasks involved in the second iteration of the sequence, drawing evidence from student submissions. Next, we show that the experiment design activity led to genuinely open-ended generation of many unique solutions. Finally we discuss positive student outcomes in self-efficacy, sense of agency, and conceptual understanding measured by surveys and a standardized conceptual learning assessment.

# DESCRIPTION OF LAB ACTIVITIES

This study took place at Cornell University in the context of MAE 3270, a junior-level mechanics of engineering materials course which is taught in the Summer and Fall terms, and is normally taught fully in person. In previous years, the course included four traditional in-person lab activities conducted in groups of 2–3 students in sections of 10-15 students. The new lab sequence was introduced in Summer 2020 in response to the COVID-19 pandemic. The activities were subsequently modified and adapted for use in Fall 2020. Cornell University welcomed students back to campus in the Fall for a mix of in-person and remote classes, however the instructional lab facility for MAE 3270 remained closed due to inability to comply with density and ventilation requirements.

In Summer 2020, MAE 3270 was taught by the first author with a total enrollment of 21 students. All class meetings were held as synchronous virtual meetings with captioned recordings available to students. A small kit was mailed to each student with material specimens, a rudimentary force gauge, a metal ruler, and assorted small hardware. Some logistical difficulties were encountered with the kits including one loss in transit overseas.



Торіс	Fall 2019	Summer 2020	Fall 2020
Lab 0: Combined loading	Traditional in-person		
Lab 1: Heat Treatment of Metals	Traditional in-person	Traditional virtual	Traditional virtual
Lab 2: Uniaxial Tension Testing	Traditional in-person	Guided home experiment, Peer-teaching video	Traditional virtual, Peer-teaching video
Lab 3: Fracture Toughness Testing	Traditional in-person	Peer-teaching video	Traditional virtual, Peer-teaching video
Lab 4: Experiment design		Young's modulus of steel wire	Young's modulus of aluminum can

In Fall 2020, MAE 3270 was taught by an experienced instructor and original creator of the course, with a total enrollment of 132 students. All lecture and lab section meetings were held as synchronous virtual meetings with captioned recordings available to students. Some discussion sections met in person while others met virtually. 95% of students were located in the same time zone as the institution or within 3 hours.

The lab activities are shown in Table 1. One of the in-person labs (combined loading) was discarded entirely. The heat treatment, uniaxial tension testing, and fracture toughness testing labs were modified for the virtual format and supplemented with peer-teaching video assignments, and a new experimental design lab was added. In the summer, the uniaxial tension testing lab was replaced with a guided home experiment on a nylon filament.

#### "Traditional" Virtual Labs

We redesigned Labs 1,2, and 3 for the remote instruction format. Students were instructed to watch a 15-20 minute pre-recorded video and complete a brief quiz before their synchronous online lab meeting. The video included a brief introduction to the lab by the instructor followed by a demonstration of the equipment and experiment with voiceover narration. The quiz tested students on basic comprehension of the relevant ASTM standard and asked them to make qualitative and quantitative predictions about the experimental results. During the online synchronous lab meeting, lab teams met separately using Zoom breakout rooms to answer a series of discussion questions, then participated in a whole-class discussion and asked questions about the lab analysis tasks. After the lab session, each team had one week to complete and submit their lab report. Teams were given a report template for each lab to provide a standard format and to give examples of good writing practices. Each successive template included less pre-written content.

#### **Peer-Teaching Video Activities**

Alongside Labs 2 and 3, students were asked to create and share 5-minute videos about mechanical testing. For the second lab, students recorded a low-fidelity demonstration of the uniaxial tension



test with household materials, describing the important aspects of specimen design, instrumentation, and validity criteria. For the third lab, students chose two materials to perform a three-point bending fracture "test" on and compared the morphology of the fracture surfaces. Students chose whimsical materials like cheese and chocolate for their comparisons. The purpose of the peer-teaching videos was to give students latitude to make independent decisions about testing and gain experience building test fixtures with household materials, but without collecting quantitative data.

# **Experimental Design Lab**

In Lab 4, students were asked to measure the Young's modulus of a metal specimen: a steel wire (provided) in Summer 2020 and an aluminum can in Fall 2020. The details that follow relate to the Fall 2020 version of the activity. The activity was presented as a real-world challenge: teams were asked to evaluate the suitability of a "material sample" (an aluminum beverage can) sent by a prospective supplier and finalize the design of a structural boom for a robotic arm for space applications (Figure 1) to be made of the candidate material. The functional requirements that the design had to satisfy (resonant frequency and local buckling criteria) were chosen so as to depend indirectly on the Young's modulus of the material. (The arm is referred to as the "IthaClaw" in reference both to the Canadarm and Cornell University's home in Ithaca, NY.) Students were allowed to assume known values for the density and Poisson's ratio, if needed for their calculations. To ensure the teaching assistants were confident in their ability to help students for this activity, all members of the instructional staff designed and performed their own experiments at home and discussed results and challenges as a group.

Students were given written guidance describing the difference between material stiffness and structural stiffness, using a comparison between a uniaxial tension coupon and a three-point bending specimen as an example. Students were also given information about three example experiments









Figure 2. Three examples of student-designed experiments for Lab 4: (a) Folded "accordion spring" design. (b) Torsional spring. (c) Diametral tension test.

by student teams from Summer 2020 who solved a different but related problem of measuring the Young's modulus of steel wire. The examples included two static deflection experiments (cantilever beam and helical spring) and one dynamic experiment (torsional pendulum). Three examples of student-designed experiments from Fall 2020 are shown in Figure 2.

Students were given a list of 14 guidance questions (see appendix) and asked to consider them when developing their experiments and to incorporate the answers into their experimental proposals. Each team brainstormed three initial ideas and received feedback from the teaching staff. They then chose one idea and wrote a brief (1-2 page) experiment proposal, which was then reviewed by instructional staff in meetings with each team. Staff were instructed to use the guidance questions to structure their feedback and to refrain from suggesting specific solutions not considered by the students themselves. Additional office hours specifically for help with the lab were scheduled at times convenient for local and overseas time zones.

Each team submitted a lab report describing their experiment and analysis of their results, as well as the final design of their boom. In addition to demonstrating that their design met the functional requirements (using their own calculated value for the Young's modulus), teams had to select and justify safety factors for each functional requirement based on the application, severity of potential failure, and their own estimate of the uncertainty of their measurement of the modulus.

# **COGNITIVE TASK ANALYSIS**

We analyzed the cognitive tasks involved in each lab activity during the Fall 2020 offering of the course according to the inventory of cognitive tasks involved in experimental physics research



# Table 2. Wieman's taxonomy of expert decision-making for experimental physics,

#### adapted.

Category	Description	Number of tasks
1	Establishing research goal	3
2	Defining criteria for suitable evidence	3
3	Determining feasibility of experiment	2
4	Experimental design	5
5	Construction and testing of apparatus	6
6	Analyzing data	5
7	Evaluating results	1
8	Analyzing implications if results are novel and/or unexpected and confirmed	1
9	Presenting the work	2
	Total	28

proposed by Wieman (2015). Although there is significant overlap between the list proposed by Wieman and the list of learning objectives for engineering labs developed by Feisel and Rosa (2005), the tasks identiied by Wieman are more granular and hierarchically organized.

Wieman's taxonomy comprises 28 tasks organized into 9 broad, roughly chronological categories, each comprising between one and six sub-tasks. The categories are summarized in Table 2 and a complete list of tasks is given in the appendix. We have adapted Wieman's original list slightly for clarity and to better match our specific context: we removed or combined three items and added three new items, including two items suggested by Burkholder et al. (2021).

We analyzed all the lab materials presented to students and identified cognitive tasks that were either explicitly given (done for the student), explicitly prompted, or implicitly expected or required based on the stated deliverables. Each lab is subdivided into specific activities performed by the student: pre-lab quiz and video, group discussion, and report for Lab 1; pre-lab quiz and video, group discussion, peer-teaching video, and report for Labs 2 and 3; and experiment proposal and report for Lab 4. Tasks identified multiple times in a single activity were only coded once.

A task was coded as "Given" if the result of the decision was described explicitly to students. A task was code as "Prompted" if the students were asked to make a particular decision, justify one made for them, or consider alternatives. A task was coded as "Expected" if the students were not explicitly instructed to do the task, but completion of the task was implied by the deliverables or by another requested task. Three examples of codes assigned to specific sentences in lab materials are given in Table 3.

The first author identified all cognitive tasks mentioned or implied in the lab activity materials and mapped each cognitive task to a specific sentence, video fragment, or instruction in the lab materials. The first author then wrote descriptions of the three codes described above and



Example	Source	Cognitive Task	Code
"Rather than performing a uniaxial tension test on every sample, we will use hardness as a proxy for strength"	Lab 1 intro	(4d) Develop a detailed data acquisition strategy.	Given
"Based on your calculations above, what is the closest allowable spacing (center-to-center) between indentations?"	Lab 1 quiz	(4b) Analyze relevant variables that may lead to systematic errors.	Prompted
"Based on your data, do you think that the hardness would eventually recover to its as-received value? If so, how long might it take?"	Lab 1 report	(6a) Model the data by suitable mathematical forms, including deciding which approximations are justified.	Expected

identified an example of each. The first and second authors independently coded the materials for Lab 1 and achieved 87% agreement. The first author coded the remaining materials independently, as this level of agreement was deemed satisfactory to support the reliability of the coding method.

The fraction of tasks at each guidance level and the total number of the (28) tasks represented are shown in Figure 3. As students gain experience making decisions over the course of the semester, the total number of tasks and autonomy afforded by each task increases. This analysis illustrates





the scaffolded design of our lab sequence, progressing from heavily prescribed activities towards an open-ended design challenge.

Evidence of engagement with cognitive tasks can be directly drawn from students' assignment submissions. For example, in the heat treatment lab (Lab 1), students are asked to predict whether the hardness of an aluminum 2024-T4 sample which was solution treated and quenched in the virtual experiment demonstration will ever recover its as-received hardness, based on measurements taken at 0, 1, 6, and 24 hours after quenching. The students are not guided about whether or how to extrapolate the data. This outcome requires task 6a: "Modeling the data by suitable mathematical forms, including deciding which approximations are justified." Since students are asked to answer based on their data, but not explicitly prompted to develop a quantitative model, the task was coded as "Expected." One group (Figure 4a) decided to fit the data with a function of the form  $H = a + b \exp(-t/\tau)$ , justifying their choice of model based on (1) the apparent shape of the data, and (2) the fact that hardness cannot increase to infinity. Another group (Figure 4b) noted the necessity of an asymptote, but decided to model their data with a fit of the form  $H = a + b \log(t)$ , perhaps not recognizing that this model predicts infinite hardness at infinite time. This example demonstrates that students may be making the same kinds of decisions as an expert, but with less technical skill.

By design, the final lab activity involves the largest number of cognitive tasks with the lowest level of support. The number of unique tasks involved in an activity gives only a rough idea of the level of autonomy afforded to the student. Next we will discuss a more precise measure of open-endedness and examine the actual designs that resulted from students' decisions.



(a) An exponential fit. (b) A logarithmic fit.



#### SOLUTION DIVERSITY ANALYSIS

A genuine and successful guided-inquiry activity should not only shift important decisions onto the students, but the goal itself should reasonably afford multiple successful approaches. Willner-Giwerc et al. (2020) proposed *solution diversity* as one measure of the success of an open-ended activity. A large number of unique (successful) solutions is evidence that the activity is truly open-ended and that students are engaging in genuine problem-solving and decision-making. A small number of unique solutions is evidence that either there are too many constraints on the problem, or that students lack the tools to make independent decisions.

We examined the procedure section of all 42 lab reports and coded the choices made by a team across a set of five decisions: the mode of the test (static deflection, dynamic, or buckling), the specimen geometry (e.g. cantilever beam, cylinder), the loading method (continuous or nearly-continuous weight, like water or salt; quantized weights, like coins or paperclips; and in one unique instance, a hydraulic car jack), the measurement tool (e.g. ruler, video camera), and amplification method (e.g. lever, shadow projection). Six reports were excluded from analysis due to incomplete information, leaving a total of 36 suitable for analysis.

The specific decisions to code were agreed upon collaboratively by the authors based on the fundamental components of mechanical testing: specimen geometry, boundary conditions, and measurement. Geometry encompasses both the shape of the specimen and the boundary conditions applied. Codes were assigned based on the actual experiment performed (as judged from the description, photograph, and schematic), not necessarily on what is claimed in the report. For example, if students claimed to have performed a test on a "simply-supported beam," but they placed each end of the aluminum strip between stacks of books, then the geometry was coded as "fixed-fixed beam." Different codes were assigned to choices if the two choices changed the experiment procedure in a substantive way: for example, while using rice vs. using coins as deadweights might seem like similar choices, the students using rice needed to measure the mass of a fixed volume of rice to establish a calibration, while the students using coins could rely on published standardized weights. Many minor decisions which were crucial to the success of the experiment, such as the method of sample preparation or orientation of the specimen along the can, were not coded. After reviewing some example reports together, the first author proceeded to code all the reports.

The space of solutions is presented graphically in Figure 5, where the height of each choice is proportional to the number of teams that selected it. We define a "solution" as the set of five specific choices made by the team, e.g. static, cantilever beam, loaded with coins, measured by ruler,





no amplification. Each path through Figure 5 represents a unique solution. Out of 36 reports with complete data, there was a total of 25 unique solutions, with no one solution being shared by more than three reports. The counting of solutions is quite robust against the particular decisions made when coding. Discarding the mode decision entirely results in the same number of solutions, and discarding the "amplification" decision reduces the number of solutions to 21. Even if the loading method is coded more generally (e.g. "hung deadweight" or "struck at tip"), there are 23 unique solutions With all of these methods, no unique solution was shared by more than four reports.

The solution diversity is apparent not just across reports, but also within each decision category. For example, 31 teams conducted static deflection tests (the most straightforward choice, and also most consistent with familiar mechanical tests studied in the class), but six teams measured resonant frequency of a structure, and one team even conducted a buckling load test. More telling is the surprising variation in sample geometry. The examples of previous experiments on steel wire shown to students included a cantilever beam, a torsional pendulum, and a helical spring. We observed 10 distinct sample geometries (including boundary conditions) ranging from no modification to the can at all, to intricate folded accordion springs from longitudinally-cut strips.



### STUDENT AGENCY AND SELF-EFFICACY

In Fall 2020, student attitudes were measured with a brief online survey taken shortly after submitting the tensile testing (Lab 2), fracture toughness testing (Lab 3), and experimental design lab (Lab 4) reports. The survey included five items about sense of agency and attitudes towards the lab activity, measured on a five-point Likert scale ("strongly disagree" to "strongly agree"), and eight items about self-efficacy on a five-point Likert scale ("not confident" to "very confident"). Three of the sense-of-agency items were adapted from Kalender et al. (2021). (The fourth item used by Kalender. et. al.—"I am in control of doing interesting experiments in a physics lab." as deamed inappropriate for our context due to our emphasis on guided inquiry, rather than open inquiry.) Survey items are shown in Table 4. We also added two items related to surprise and novelty. Kalender's original 4-item assessment has been validated elsewhere (Kalender et al., 2020). We calculated Cronbach's  $\alpha$  for the sense of agency items ( $\alpha = 0.72$ ) and self-efficacy items ( $\alpha = 0.80$ ).

Out of 132 students, 75 (57%) completed all three surveys. Mean endorsement level for the attitude items and mean confidence level for the self-efficacy items are shown in Figure 6. The effect of the lab was investigated using a repeated (within-subjects) ANOVA and subsequent post-hoc Tukey tests of differences between labs. Results for items with significant or marginally significant differences are given in Table 5.

Students' feeling of control over goals and analysis increased dramatically between Lab 3 and Lab 4, as expected. The fraction of students endorsing the control of goals and control of analysis items increased from 53% to 83% and from about 63% to 92%, respectively. However, there was no discernible difference for the related item "I have the freedom to create my best work for this

Sense of agency: Five-point Likert scale: strongly disagree to strongly agree. I am in control of setting the goals for this lab activity.* I am in control of choosing the appropriate analysis tools to evaluate experimental data.* I have the freedom to create my best work for this activity.*	$\alpha = 0.72$
I was sometimes surprised by the outcomes during this lab activity.	
I learned something unexpected during this lab activity.	
Self-efficacy: Five-point Likert scale: Not confident to Very confident.	$\alpha = 0.80$
Express my opinions when others disagree with me.	
Achieve an equitable division of work within my group.	
Overcome any problems I encounter during the experiment or analysis.	
Interpret data taking into account experimental uncertainty.	
Interpret graphs of experimental measurements.	
Make accurate predictions about experimental outcomes.	
Design an experiment to reliably measure mechanical properties.	
Generate further questions based on my observations in the lab.	





Table 5. Results of an Analysis of Variance (ANOVA) for each survey item with corrected pairwise comparisons. Only survey items with significant or marginally-significant differences are shown.

	ANOVA			Pa	irwise Tukey's HS	D
Item	dof	F	p	Lab $2 \rightarrow 3$	Lab $2 \rightarrow 4$	Lab $2 \rightarrow 4$ Lab $3 \rightarrow 4$
Attitudes						
Control of goals	134.0	31.7	****	ns	****	****
Control of analysis	132.5	28.3	****	ns	****	****
Surprised by outcomes	136.0	17.3	****	ns	****	**
Self-efficacy						
Design an experiment	148.0	18.7	****	ns	****	**
Make predictions	148.0	9.82	****	ns	**	*
Interpret data	148.0	2.40	0.094	ns	ns	ns



activity." Although we did not conduct validation interviews, it's possible that students are interpreting "freedom to create" as the ability to achieve a good grade or create a polished product, not as the latitude to make independent decisions which lead to success. Students also experienced more surprise about the outcomes of Lab 4, compared with Labs 2 and 3.

The survey included two items directly related to the primary learning objectives of the lab sequence (designing experiments and making predictions), three secondary generic objectives (generating further questions, interpreting data, and interpreting graphs), as well as three items related to important, but non-targeted skills (expressing opinions, distributing labor, and overcoming difficulties). Large, significant increases over time were observed for confidence in designing an experiment (Cohen's d = 0.68) and making predictions (Cohen's d = 0.49). 52% of students reported increased confidence on designing experiments, while only 9% reported a decrease. The trend with respect to time (Lab 2 < Lab 3 < Lab 4) were the same for both primary learning objectives, however in both cases the pairwise difference between Lab 2 and Lab 3 was not significant. We did not observe significant increases in self-reported confidence in generating further questions, interpreting data, or interpreting graphs. The lack of a similar trend in the non-targeted skills suggests that the increase in the targeted skills is due to the lab intervention, and not to differing survey response patterns over time.

#### LEARNING OUTCOMES

We previously developed a multiple-choice conceptual learning assessment for mechanics of materials tailored to the learning outcomes of MAE 3270 (Ford et al., 2020). The assessment has been validated with think-aloud interviews with students, review by experts, item analysis, and concurrent validity with existing metrics. The assessment was completed by students online through the course learning management system. In Fall 2019 (n = 96) the assessment was divided into three quizzes according to subject area and delivered before each major exam. In Fall 2020 (n = 121) the entire assessment was completed at the end of the term. (Summer 2020 had a small sample size, n = 22, and was not included in this analysis.)

Eight items were identified from the entire assessment which deal with concepts aligned with the lab outcomes including material processing (Q36), uniaxial tension testing and viscoelasticity (Q32, Q33, and Q34), fracture toughness (Q21 and Q23), combined loading (Q1), and strain gauge measurements (Q17). Assessment scores and individual item scores are presented in Figure 7. The lab activity on combined loading and strain gauge measurements (Lab O) addressing the concepts in Q1 and Q17 was removed in Fall 2020. Results are also given in tabular form in the appendix.





Results between years were compared using an un-paired t-test (for average scores on multiple items) or a proportion test (for the fraction of students answering correctly on a single item). Overall assessment scores were very similar between Fall 2019 and Fall 2020 (t-test: p = 0.816). Average scores on only the items associated with lab topics increased slightly in Fall 2020, but the increase was not significant (Cohen's d = 0.18, t-test: p = 0.19), suggesting that conceptual learning did not suffer as a result of removing the in-person lab component. Indeed, students even performed slightly (Q1: +8 percentage points, p = 0.29) and significantly (Q17: +24 percentage points, p < 0.001) better on the combined loading and strain gauge items.

Two significant effects should be considered when comparing the 2019 and 2020 results: First, the Fall 2020 students were likely under considerably more extra-curricular stress due to the COVID-19 pandemic (Copeland et al., 2021; Salari et al., 2020). Second, in Fall 2019 students completed the assessment in three installments, with a shorter interval between learning the relevant concepts and taking the test. The redesign of the lab activities was the most significant part of a broader project which included the development of some in-class activities and other efforts to mitigate the negative effects of remote instruction. In light of these considerations, we argue that the small and statistically insignificant improvement in scores in Fall 2020 is actually a conservative estimate of the improvement in conceptual learning due to these efforts.



# DISCUSSION

A primary goal of engineering education is to help students develop an expert-like mindset that helps them transfer acquired knowledge to new problems and generate new knowledge when necessary, a set of skills and habits called "interpretive knowing" (Etkina et al., 2010). Supporting students' decision-making agency helps them take ownership over the process and results of the activity to develop interpretive knowing. As a specific example, increased ownership over process and product may cause students to actively reflect on and process experimental results in the face of statistical uncertainty. In an earlier iteration of the experiment design lab in Summer 2020, students were tasked with measuring the Young's modulus of a steel wire, which was mailed to each student. The instructor measured the wire diameter for the students to use in their analysis. One student, puzzled by their experimental result, which was out of the typical range for steel even after estimating experimental uncertainty, made more careful measurements and discovered that the wire cross-section was elliptical after an extensive troubleshooting process. Another student discovered the same issue when, during the course of their experiment which involved torsion, the wire surface developed a clearly noticeable helical "fluting" pattern. Both of these observations were made outside of the usual "measure and report" requirements of a traditional lab, and required calling into question the given information. In our experience as instructors we have seen many examples of students blindly reporting results that should have raised red flags, and sweeping any discrepancies with expectations into the dust bin of "experimenter error."

Although the redesigned lab activities were partially motivated by the COVID-19 pandemic, they also improve on the traditional in-person offerings in several ways. The new lab sequence affords students the opportunity to practice making and executing design decisions oriented towards a tangible goal. Home experiments make use of cheaper materials and smaller forces, and the risks to personal safety or equipment are minimized. The experiment design activity was designed with equity in mind: aluminum beverage cans are cheap, available globally, and can be cut with house-hold scissors. Many successful experiments used nothing more sophisticated than a ruler. They help students appreciate the engineering principles in the context of their everyday environment. The unavoidable limitations of performing experiments without expensive equipment open up a broad space of potential successful solutions. In these regards, it is superior to the type of recipe-based labs that we used previously when teaching in person. Next, we will discuss some drawbacks to the inquiry-based lab and suggest potential improvements.

The cost of materials was minimal, but the redesigned lab sequence required some additional staff time beyond what would have been required for in-person lab activities. The teaching assistants expressed concern that they didn't have enough experience with activities like the experimental

design lab to give useful feedback to students. In order to gain experience and anticipate student concerns, the teaching staff brainstormed several possible approaches to the experiment, and then each member conducted their own experiment at home, making sure to choose a variety of experiment types. Afterwards, the first author led a discussion about potential challenges, measurement precision, and practical realization of idealized boundary conditions. This process alleviated the teaching assistants' concerns and helped prepare them to assist students.

Student satisfaction with the labs in Fall 2020 was mixed: 72% of students rated the value of the laboratory activities as moderately to very valuable, but the majority of free-response comments were either resigned (e.g. "I didn't like having an online lab but I feel like the course staff did the best with what they had to work with.") or negative (e.g. "I did not feel that having us perform the labs at home was helpful for my understanding of the concepts.") Most students feel they missed out on something essential by not being able to operate the test machinery themselves.

How might the benefits of this activity be adapted to in-person instruction to give students a richer and more satisfying experience? First, the final goal—for students to develop and analyze a novel experiment—should remain the same. The only way to teach experimental design is to have students design experiments. Lab facilities could be equipped with more general-use tools, such as calipers and small force gauges, that students could choose to incorporate into their experiment. The design lab could be scaffolded by having students develop and test a portion of the experiment (e.g. design, build, and calibrate their own force sensor) that will be used for their final experiment. Finally, the collaborative element could be enhanced by turning the activity into a competition: e.g. the students wouldn't know the exact form of the specimen until their lab day, during which they would need to use the tools they developed in previous activities to measure the property of interest.

# ACKNOWLEDGMENTS

The authors would like to thank Prof. Alan Zehnder for his enthusiastic support for this project, as well as Prof. Natasha Holmes and Dr. Z. Yasemin Kalender for inspiring discussions about skillsbased laboratory instruction. This paper is based upon work supported by the Cornell Active Learning Initiative. This study was approved by the Cornell University Institutional Review Board under protocol #1708007347.

All statistical analysis was performed in R with the tidyverse package (R Core Team, 2021; Wickham et al., 2019). Statistical analysis was performed using the rstatix package (Kassambara, 2021). Plots were generated using the ggplot2 (Wickham, 2016) and ggalluvial packages (Brunson and Read, 2020) and post-processed using Inkscape (Inkscape Project, 2021).



#### REFERENCES

ABET (2018). 2019-2020 Criteria for Accrediting Engineering Programs.

Ankeny, C. J. and M. C. Tresch (2021). Creation and Deployment of a Virtual, Inquiry-Guided Biomedical Engineering Laboratory Course. *Biomedical Engineering Education 1*, 67–71.

Banchi, H. and R. Bell (2008). The many levels of inquiry: inquiry comes in various forms. *Science and Children 46*(2), 26–29. Benedict, B. S., K. R. Maxey, D. Verdin, and A. Godwin (2020). A Review of Agentic Frameworks in Engineering Education. In *2020 ASEE Virtual Annual Conference Content Access*.

Brunson, J. C. and Q. D. Read (2020). ggalluvial: Alluvial plots in ggplot2. R package version 0.12.3.

Burkholder, E., L. Hwang, E. Sattely, and N. Holmes (2021). Supporting decision-making in upper-level chemical engineering laboratories. *Education for Chemical Engineers* 35, 69–80.

Copeland, W. E., E. McGinnis, Y. Bai, Z. Adams, H. Nardone, V. Devadanam, J. Rettew, and J. J. Hudziak (2021, January). Impact of COVID-19 Pandemic on College Student Mental Health and Wellness. *Journal of the American Academy of Child & Adolescent Psychiatry* 60(1), 134–141.

Etkina, E., A. Karelina, M. Ruibal-Villasenor, D. Rosengrant, R. Jordan, and C. E. Hmelo-Silver (2010, January). Design and Reflection Help Students Develop Scientific Abilities: Learning in Introductory Physics Laboratories. *Journal of the Learning Sciences* 19(1), 54–98.

Feisel, L. D. and A. J. Rosa (2005). The Role of the Laboratory in Undergraduate Engineering Education. *Journal of Engineering Education* 94(1), 121-130.

Ford, M., H. Ritz, B. Finio, and E. Fisher (2020). Multiple-choice Learning Assessments for Intermediate Mechanical Engineering Courses: Insights from Think-aloud Interviews. In *2020 ASEE Virtual Annual Conference Content Access*, pp. 34990.

Fox, M. F., A. Werth, J. R. Hoehn, and H. J. Lewandowski (2020). Teaching labs during a pandemic: Lessons from Spring 2020 and an outlook for the future. arXiv: 2007.01271.

Halstead, E. (2016). Effects of reducing scaffolding in an undergraduate electronics lab. *American Journal of Physics* 84(7), 552–556.

Hartocollis, A. (2021, February). U.S. Universities Plan for a 'More Normal' Fall. The New York Times.

Holmes, N. G., J. Day, A. H. Park, D. A. Bonn, and I. Roll (2014). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instructional Science* 42(4), 523–538. Holmes, N. G., B. Keep, and C. Wieman (2020). Developing scientific decision making by structuring and supporting

student agency. *Physical Review Physics Education Research 16*, 010109.

Holmes, N. G., J. Olsen, J. L. Thomas, and C. E. Wieman (2017). Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content. *Physical Review Physics Education Research 13*, 010129.
Holmes, N. G. and E. M. Smith (2019). Operationalizing the AAPT Learning Goals for the Lab. *The Physics Teacher 57*(5), 296–299.
Holmes, N. G. and C. E. Wieman (2016). Examining and contrasting the cognitive activities engaged in undergraduate

research experiences and lab courses. Physical Review Physics Education Research 12, 020103.

Inkscape Project (2021). Inkscape. version 1.1.1.

Jaschik, S. (2021, August). Delta Variant Raises Questions as Campuses Start Semester. Inside Higher Ed.

Johnson, B. E. and J. W. Morphew (2016). An analysis of recipe-based instruction in an introductory fluid mechanics laboratory. In *ASEE Annual Conference & Exposition*.

Kalender, Z. Y., M. M. Stein, and N. G. Holmes (2020, September). Sense of agency, gender, and students' perception in open-ended physics labs. In *2020 Physics Education Research Conference Proceedings*, Virtual Conference, pp. 259–265. American Association of Physics Teachers.



Kalender, Z. Y., E. Stump, K. Hubenig, and N. Holmes (2021, October). Restructuring physics labs to cultivate sense of student agency. *Physical Review Physics Education Research 17*(2), 020128.

Kassambara, A. (2021). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.0.

Lazonder, A. W. and R. Harmsen (2016, September). Meta-Analysis of Inquiry-Based Learning: Effects of Guidance. *Review of Educational Research 86*(3), 681-718.

Morrison, A. (2014). From cookbooks to single sentences: The evolution of my labs. *The Physics Teacher 52*(8), 505–506. O'Neill, D. P. (2021). Redesign of a BME Lab Class to Maintain Hands-on Experimentation Despite Remote Learning Constraints. *Biomedical Engineering Education 1*, 229–235.

R Core Team (2021). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Ritz, H., M. N. Silberstein, and N. Andarawis-Puri (2018). Uniaxial tension testing lab: Fewer instructions for better results? In ASEE Annual Conference & Exposition.

Salari, N., A. Hosseinian-Far, R. Jalali, A. Vaisi-Raygani, S. Rasoulpoor, M. Mohammadi, S. Rasoulpoor, and B. Khaledi-Paveh (2020, December). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Globalization and Health* 16(1), 57.

Ural, E. (2016, February). The Effect of Guided-Inquiry Laboratory Experiments on Science Education Students' Chemistry Laboratory Attitudes, Anxiety and Achievement. *Journal of Education and Training Studies* 4(4), 217–227.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software* 4(43), 1686.

Wieman, C. (2015). Comparative Cognitive Task Analyses of Experimental Science and Instructional Laboratory Courses. *The Physics Teacher* 53(6), 349–351.

Wilcox, B. R. and H. J. Lewandowski (2016). Open-ended versus guided laboratory activities: Impact on students' beliefs about experimental physics. *Physical Review Physics Education Research 12*, 020132.

Willner-Giwerc, S., K. B. Wendell, C. B. Rogers, E. E. Danahy, and I. Stuopis (2020). Solution diversity in engineering computing final projects. In 2020 ASEE Virtual Annual Conference Content Access.

#### **AUTHORS**



**Matthew J. Ford** (he/him) received his B.S. 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 a postdoc with the Cornell Active Learning Initiative, he joined the School of Engineering and Technology at UW Tacoma to help establish its new mechanical engineering program. His teaching and research interests include solid mechanics, engineering design, and inquiry-guided learning. He has supervised undergraduate



and master's student research projects and capstone design teams. He is a member of the American Society for Engineering Education.



**Soheil Fatehiboroujeni** (he/him) received his Ph.D. in Mechanical Engineering from the University of California, Merced in 2018 focused on the nonlinear dynamics of biological filaments. As an engineering educator and postdoctoral researcher at Sibley School of Mechanical and Aerospace Engineering at Cornell University, Soheil worked in the Active Learning Initiative (ALI) to promote student-centered learning and the use of computational tools such as Matlab and ANSYS in authentic engineering contexts. In the Spring of 2022 Soheil joined the Colorado State University as an Assistant Professor of Practice in the department

of mechanical engineering. His research is currently focused on the long-term retention of knowledge and skills in engineering education.



**Elizabeth M. Fisher** (she/her) received her Bachelor's degree in physics from Yale and her M.S. and Ph.D. degrees in Engineering Science: Mechanical Engineering from the University of California, Berkeley. She is an Associate Professor in the Sibley School of Mechanical and Aerospace Engineering at Cornell University. Her research focuses on combustion chemistry and thermal transformations of biomass. She teaches thermal sciences, combustion, and propulsion courses. She has been a visiting researcher at the National Renewable Energy Laboratory, Leeds University, Lawrence Livermore National Laboratory, and the Naval Research Laboratory.



**Hadas Ritz** (she/her) 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 2021 ASEE National Outstanding Teaching Award.



# APPENDIX

## **Guidance Questions**

The following guidance questions were given to students to help them structure their experiment proposals. The teaching staff used these questions as a template for giving feedback to teams.

- 1. Test specimen design
  - (a) How will you make the specimen? Can the method be easily replicated by all your group members?
  - (b) What geometric parameters do you need to control and measure?
  - (c) How will you calculate the stiffness of your structure?
- 2. Experimental design
  - (a) How are the boundary conditions (loads and supports) applied to your specimen? Are your assumptions justified?
  - (b) What maximum load do you expect to apply? Under this load, what maximum stress do you expect?
  - (c) Under this load, what maximum deflection do you expect? Are the assumptions of your mechanical analysis still reasonably valid after this displacement? How could you verify that?
- 3. Measurement and uncertainty
  - (a) What measurements will you make, and how?
  - (b) How can you estimate the uncertainty of each of these measurements?
  - (c) Will you measure displacement? If so, what is the smallest displacement you could reasonably measure, and how does it compare to your expected maximum displacement?
  - (d) Will you measure or apply forces? If so, what is the smallest force you could reasonably measure, and how does it compare to your expected maximum applied force?
  - (e) How will you calibrate your applied force? (How will you know what force you are applying, or how will you know that whatever you use to measure forces is accurate?
- 4. Troubleshooting
  - (a) Have you tested out your experiment, even very roughly?
  - (b) What difficulties did you run into? What are some potential concerns?
  - (c) What simple design changes could you make, and what performance tradeoffs would result?

#### Wieman's Cognitive Tasks

List adapted from Wieman (2015) and Burkholder et al. (2021). The task labels are consistent with Wieman (2015). Added or modified items have been noted as such. Ellipses (...) indicate truncation of a longer description by Wieman.



- 1. Establishing research goal
  - (a) Deciding if the goal is interesting, timely, worthwhile.
  - (b) Predicting if the goal is sufficiently ahead of current knowledge.
  - (c) Evaluating whether the research question is consistent with the constraints on funding, time, equipment, and laboratory capacity, including personnel.
- 2. Defining criteria for suitable evidence
  - (a) What data would be convincing given the state of the field?
  - (b) What variables are important and how might they be measured and controlled?
  - (c) What types of experimental controls and checks would need to be in place?
- 3. Determining feasibility of experiment
  - (a) Predicting whether or not experiment is realistically possible.
  - (b) The researcher must also analyze contingency options.
- 4. Experimental design
  - (a) Exploration of many possible preliminary designs.
  - (b) Analyzing relevant variables that may lead to systematic errors.
  - (c) Finalizing the design...
  - (d) Developing detailed data acquisition strategy...
  - (e) (added) Developing predictions for the experimental results.
- 5. Construction and testing of apparatus
  - (a) (modified) Building or assembling the apparatus
  - (b) Developing criteria and test procedures for evaluation of the individual apparatus components.
  - (c) Collecting data on performance of specific components and full apparatus.
  - (d) Developing procedures for tracking down the source of malfunction...
  - (e) Figuring how to modify particular parts...
  - (g) Collecting experimental data.
- 6. Analyzing data
  - (a) Modelling the data by suitable mathematical forms, including deciding which approximations are justified.
  - (b) Deciding on what analysis methods are procedures are appropriate.
  - (c) Calculating the statistical uncertainty.
  - (d) Calculating the systematic uncertainty.
  - (e) (added) Interpreting the data relative to the model.
- 7. Evaluating results
  - (b) Testing data that come out as expected



- 8. Analyzing implications if results are novel and/or unexpected and confirmed
  - (a) What are plausible interpretations...
- 9. Presenting the work
  - (a) Follow standard data display procedures or, as needed, develop new procedures that highlight critical features of methods or results.
  - (b) Explain the work so the broader context and uniqueness of the work, the apparatus, the procedures, and the conclusions are easily understood...

# **Conceptual Assessment Items**

Learning assessment results are given in Table 6.

The conceptual assessment items dealing with lab-related concepts are shown below. Each item has a question number (e.g. Q1) and an item identifier (e.g. SA-SS-01). Items are organized by lab topic.

	Averag		
Learning assessment items	Fall 2019 ( <i>n</i> = 96)	Fall 2020 ( <i>n</i> = 121)	р
Combined loading and strain gauges			·
Q1	53%	61%	0.29
Q17	22%	45%	0.0005
Processing			
Q36	33%	36%	0.85
Tensile testing			
Q32	74%	75%	0.95
Q33	35%	38%	0.80
Q34	54%	52%	0.86
Fracture			
Q21	82%	81%	0.94
Q23	58%	56%	0.86
All lab-related items (8)	51.6%	55.6%	0.19
All items (37)	50.9%	51.4%	0.82

Average score for individual items is the proportion of students who answered correctly. *p*-values were calculated in R using a proportion test (for single items) or a t-test (for averages).





- (b) Top 90° and inclined  $45^{\circ}$
- (c) Top  $0^{\circ}$  and middle  $90^{\circ}$
- (d) Middle 0° and inclined 45°
- (e) Top  $0^{\circ}$  and middle  $45^{\circ}$





A metal workpiece is subjected to a series of unknown processing steps. Which of the following properties is *least* likely to change as a result of the processing?

- (a) Yield strength
- (b) Shear modulus
- (c) Ductility
- (d) Fracture toughness
- (e) They are all equally likely to change.

Box 3. Q36 (MB-PR-02): Material processing.







A viscoelastic specimen is quickly stretched to a certain length at  $t = t_0$  and then held for a long time. The stress slowly decreases over time. stress time  $t_0$  $t_1$ At  $t - t_{i}$ , the specimen is quickly returned to its original length. What is the most likely graph of stress vs. time? stress stress stress stress time time  $t_1$  $t_0$  $t_1$  $t_0$  $t_0$  $t_0$ (b) (c) (a) (d) Box 6. Q34 (MB-CL-O3): Stress reversal in viscoelastic specimen.





