WHAT'S HAPPENING IN LAB?

MULTI-DIMENSIONAL ASSESSMENT TOOLS TO TRACK STUDENT EXPERIENCE THROUGH A UNIT OPERATIONS LABORATORY SEQUENCE

HEATHER C.S. CHENETTE, GREGORY T. NEUMANN, AND DANIEL D. ANASTASIO

Rose-Hulman Institute of Technology • Terre Haute, IN 47803

MOTIVATION

"Lab was basically just the same thing over and over. I don't think I actually learned anything after that first experiment. It was just the same thing on a new piece of equipment. The department really needs to overhaul this course into something worthwhile."

Student evaluation comments like these are likely familiar to chemical engineering laboratory instructors. Sure, on average the students in your class seem to be collecting adequate data and writing effective reports, but students admitting to a perceived lack of learning would give most educators pause.

Laboratories are a key component of ABET-accredited chemical engineering programs. However, many chemical engineering educators express reservations about teaching the laboratory course. The amount of essential and unique learning opportunities that must be conveyed within laboratory settings can be daunting. The major graded assignments are written reports that are very time-consuming to grade, and many times students neglect to read and address the detailed feedback they are given. Particularly vocal students may be inclined to share their opinions with laboratory instructors, often complaining that the structure of a course or the experiments are boring, dated, irrelevant, tedious, repetitive, time-consuming, or doomed to failure. Even if these comments come from a small handful of students, it can be difficult to filter out this noise when judging the current state of a laboratory course. This combination of factors can be demoralizing to a laboratory instructor, making that instructor wonder what, if anything, students are taking away from the laboratory course and whether sweeping changes should be introduced to the laboratory curriculum. These large decisions should realistically be made with as much data as possible, so it is useful to gather more information to see if these individual comments have any merit.

In laboratory courses students gain vital experience required for common chemical engineering jobs and learn material that is difficult to teach through a traditional classroom experience. While engineering laboratories may have lacked coherent or unified learning objectives in the past, efforts have been made to identify common laboratory objectives regardless of discipline.[1] Feisel and Rosa present a list of

Heather C.S. Chenette, *PhD, is an assistant professor of chemical engineering at Rose-Hulman Institute of Technology. She earned her PhD from Clemson University. She has six years of experience teaching unit operations laboratory courses. Her research interests include membrane separations in biopharmaceutical applications and exploring learning experiences of undergraduate engineering students through qualitative and quantitative methods.*

Gregory T. Neumann, *PhD, is an assistant professor of chemical engineering at Rose-Hulman Institute of Technology. He earned his PhD from the University of Notre Dame. He has taught unit operations laboratory for the past 5 years and his research interests focus on lab assessment improvements and development of heterogeneous catalysts in an education setting.*
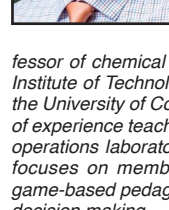
Daniel D. Anastasio, *PhD, is an assistant professor of chemical engineering at Rose-Hulman Institute of Technology. He earned his PhD from the University of Connecticut. He has eight years of experience teaching chemical engineering unit operations laboratory courses, and his research focuses on membrane separations and using game-based pedagogies to teach process safety decision making.*

thirteen fundamental objectives encompassing topics such as Design, Communication, and Safety. A recent survey of chemical engineering departments shows the most common learning outcomes for laboratory courses; it also reveals that educators do not assess each of these outcomes directly,[2] which is problematic. For example, consider a typical course objective theme: *to become familiar with appropriate instrumentation*. Recognizing that "becoming familiar" is not an observable course objective, this theme may manifest as an ability to "read", "choose", "troubleshoot", or "identify" individual pieces of instrumentation. How does an educator know whether students are in fact *familiar with appropriate instrumentation* if it is not assessed directly? As educators reflect on the effectiveness of their teaching, the curriculum, and the proposed changes in a laboratory course, how does one know if those changes maintain or improve student learning? How many credit hours in lab are actually needed to achieve these outcomes?

This study sought to close assessment gaps by using multiple assessment instruments to capture both the broad and specific student experience in the unit operations laboratory. While the benefits and possibilities of coupled direct and indirect assessment have appeared in engineering education literature,[3] this type of assessment is more prevalent in disciplines such as the performing arts and social work.[4-7]

To design the study, a conscious effort was made to employ appropriate assessment instruments. The authors began with acknowledging the student learning experience is made up of their attitudes toward, their knowledge of, and their abilities in a subject. We also acknowledged that the learner's perception may be at odds with a direct assessment of those attributes. The design of the study brings together multiple assessments to gain access to various dimensions of the student learning experience as shown in the framework in Table 1.

The MUSIC Model Inventory[8] aligned well with our desire to capture self-evaluated attitudes related to the laboratory course. The Undergraduate Research Student Self-Assessment (URSSA),[9] with slight modifications, overlapped with our intentions to measure self-assessed knowledge and abilities of laboratory skills. This study did not embark to collect direct assessments of student attitudes from peers or instructors. However, to directly assess students' knowledge and abilities of select laboratory skills, a Skills Test was developed. An existing evaluation tool, the VALUE Rubric for Written Communication, was employed to measure students'

abilities in this domain. These instruments will be described in greater detail in the next section.

Together the instruments chosen or developed for this study assessed multiple dimensions of the student laboratory experience, capturing both students' self-perceived ability and their actual skills and knowledge relating to a wide range of laboratory learning objectives. We acknowledge that student self-assessment of ability or skills can be limited.[10] However, insights from this exercise can be useful as educators seek to know how students perceive their own learning compared to the direct assessment of knowledge and skills, as departments desire routine evaluations within the laboratory, and as educators need evidence to make data-driven changes to a laboratory curriculum. We will discuss the instruments used in this study, explain why and how the assessments were administered, and discuss the resulting data collected from senior chemical engineering students taking the chemical engineering laboratory course at Rose-Hulman Institute of Technology (RHIT).

INSTRUMENT BACKGROUND

Student Self-Assessment Instruments

For the purposes of instrument reliability, two existing student self-assessment instruments were employed in this study. Each instrument was modified slightly to better align with chemical engineering laboratory courses. Modifications are described in greater detail in the next section.

One self-assessment instrument selected was the MUSIC Model of Academic Motivation Inventory. The MUSIC Model Inventory makes use of a conceptual framework with five categories (as per the acronym): eMpowerment, Usefulness, Success, Interest, and Caring. It is rooted in educational psychology research on student motivation.[8, 11, 12] The validity of this instrument for use with undergraduate students was established via confirmatory factor analysis and generated Pearson's correlation coefficients.[12] To speak to the success of the tool in other experiential learning contexts, versions of the instrument have been applied and tested for validity evidence in several fields and educational settings ranging from elementary school music students[12] to university students in summer bridge and first-year engineering programs[13, 14] to pharmacy students in required pharmacy courses.[15] In a recent study of 355 students in a psychology course, researchers coupled this instrument with other assessments and found that students' "*class perceptions predicted their engagement, which then predicted their learning*."[16]

While the study pertains to a different discipline, this is no doubt a profound result that gives us confidence in the quality

TABLE 1
Framework for capturing multiple dimensions of the student experience.

Perspective / Experience	Attitudes	Knowledge	Abilities
Assessed by self (student)	MUSIC	URSSA	URSSA
Assessed directly (evaluator)	–	Skills Test	Skills Test, VALUE

of the assessment as a whole. While the present study does not aim to make predictions or perform rigorous statistical analysis about how learning and engagement relate, we found this tool relevant and reliable for our aim to quantify student motivation and how it changes with time in the context of a laboratory course. Furthermore, while there may be aspects of this study, such as fewer students in the cohort and variability in the project equipment from student to student, that limit its success, the tool is flexible in its administration and use, supported by developers' statement that "*there is no one 'correct' way to use the inventory scores*."[12]

The second self-assessment employed was a modified version of the Undergraduate Research Student Self-Assessment (URSSA). The URSSA was developed to specifically assess student perceptions of their learning in research experiences and programs within life sciences[9] and is based on the result of a longitudinal study involving student interviews.[17, 18] Some aspects of the statistical validity have been explored indicating mixed results from various factor models and high reliability.[18] A criticism of the tool is that it lacks a theoretical framework, thus limiting its use to curating lists of "best practices" instead of providing a complete understanding of student learning that can be related across different educational programs.[19] However, the aim of this work aligns with the intended use of the instrument as a diagnostic tool for groups of ten or more students. Additionally, many aspects of undergraduate research learning experiences captured by the URSSA overlap with elements of unit operations laboratories, making this a suitable match for this study. Moreover, the URSSA is an adaptable instrument. In the directions for use, administrators are given an opportunity to delete, move, or edit the wording of the questions until they are satisfied with their own URSSA version.[20] It should be noted that, although survey administrators are encouraged to alter the survey for their benefit, the validity of the instrument only applies to the original URSSA items. The modifications made to the original URSSA are described in the next section.

Direct Assessment Instruments

Direct assessment of student knowledge and skills related to laboratory learning objectives required the development of a laboratory skills test (LST).[21] The test's questions were modeled after concept inventory questions like those in the AIChE Concept Warehouse.[22] Question topics were drawn from laboratory planning discussions among RHIT chemical engineering faculty and included emphasis on safety, ability to acquire and analyze data, linking experimental results back to theory, the troubleshooting of both equipment and data analysis methods, knowledge of equipment and instrumentation, time- and team-management skills, application of new knowledge, and written and verbal communication. Ultimately twenty-two questions were produced for the skills test

used in this assessment based on the major laboratory topics. A known limitation is that not all learning objective topics were adequately captured by the concept inventory format of the skills test. Other RHIT chemical engineering faculty reviewed and verified the question content and answers. Question formats for this skills test included multiple choice with one or more correct response, matching, labeling, and numerical data entry. Each laboratory objective was assessed across multiple questions.

While the LST was designed to measure student mastery of concepts related to the safe and reliable acquisition and analysis of experimental data, few questions focus on the communication of results. Therefore, to directly assess students' communication skills, the Written Communication VALUE Rubric[23] was applied to written student work. This rubric was designed to provide guidelines to assess written student work for context, content development, conventions, sources, and writing mechanics. These general writing skills aligned well with the needs of a laboratory report, were not specific to any one experiment, and could be easily assessed by any faculty member, even those without chemical engineering expertise.

EXPERIMENTAL METHODS

Laboratory Sequence Description

Students typically enter the three-course lab sequence during the spring quarter of their penultimate year in the curriculum and proceed during the fall and winter quarters of their final year. During each course, the student is assigned to a three-person team and one instructor (instructors advise 2-3 teams at a time). In the first course students meet for one three-hour session per week, and their project focus is one experimental apparatus. Separate instruction and assignments exist to support students during their in-lab experience. During the second and third courses in the sequence, students meet for two three-hour sessions per week, and they explore two projects in each ten-week course. Repeat projects, instructors, and team members are avoided via intentional instructor assignment. For each project students engage in some level of project planning, data collection and analysis, and individual report writing, although the scope of these elements varies by course. Each written report goes through a series of revisions, which consist of a non-graded peer-reviewed draft, an initial report submitted for grading, and a final report revised based on instructor feedback to the graded initial report. At the end of each ten-week course, students present their findings from one project in the form of an oral report given to lab instructors and their student peers. Thus, by the end of the three-course sequence, students have engaged with five projects, worked with three different sets of teams and instructors, given three oral presentations, and written five documents of varying length, style, and focus.

Self-Assessment and Skills Test for Assessing Student Abilities and Knowledge

The self-assessment instruments (MUSIC Model Inventory and URSSA) and LST were administered to juniors during the first week of their first chemical engineering laboratory course (early March). The same cohort of students, as seniors, took the self-assessment and skills test again during the last week of their final chemical engineering laboratory course in a three-course sequence (late February the following calendar year). Only minor changes to limited questions were made to provide some variety between the two tests. To assure impartiality of the investigators, who may have been acting as instructors for the laboratory course, the surveys were administered by a separate institutional assessment office. Data were held by this office until the entire laboratory sequence was complete, and all personal identifiers were removed prior to analysis by the investigators. For the written communication assessment, the eight-page written project reports were collected during the second and third laboratory course for external evaluation. Only students who completed all three courses in sequence were included in this study (N = 58).

To reveal overarching trends about student *attitudes* between the start and end of the complete laboratory sequence, the student rankings from the adapted MUSIC Model Inventory (1 = lowest and 6 = highest) were compiled into averages and standard deviations across the entire cohort for the four separate metrics of eMpowerment, Usefulness, Success, and Interest. Because the purpose of this study was to assess student engagement with the course, not the individual course instructors, the Caring vector was excluded, yielding a twenty-item inventory. In addition to course averages, the distributions of rankings within these four aspects were also analyzed for both the initial and final survey responses.

To illuminate changes in students' *self-perceived knowledge* between the start to the end of the laboratory sequence, we analyzed quantitative Likert-scale data collected from a survey modeled after the URSSA. Some topics found in chemical engineering laboratories are not present in the most recent version of the URSSA, such as teaming and safety. Therefore, twelve additional questions were added relating to these topic areas following a template provided by the creators of the instrument to assure appropriate question prompts. Additionally, it was our aim to track changes in student self-assessment over time. Hence, the verb tense was modified to align with the timing of the survey. Ultimately, we administered a revised URSSA survey consisting of thirty-six Likert scale questions focusing on students' own perceived abilities related to laboratory knowledge and skills.

To observe how students' *demonstrated knowledge* changed, we analyzed quantitative performance data from the LST. To quantify changes for each student in the study, an individual student "distance traveled" or "movement" score was calculated for each question on the URSSA and LST. For this study, the movement of a student's response was calculated as follows, modelled after the calculation of a normalized gain score:

$$movement = \frac{final - initial}{maximum - initial} \qquad (1)$$

where the *maximum* values were the highest possible ranking on the URSSA questions (ranking of 6) or a perfect answer on the LST question (1 point). Four ranges of movement were established to interpret shifts in student performance and experience. Those that shifted toward higher rankings or higher scores were categorized as a positive change, identified in subsequent figures as "POS", and those that shifted lower were categorized as a negative change, or "NEG." Students whose response to a specific question showed no change were categorized as "ZERO." However, students who ranked or scored the maximum value on their initial assessment and also exhibited a maximum value on the final assessments were placed in a separate "MAX" category and omitted from the "ZERO" category. The percentage of the entire cohort that fell into each movement category (POS, NEG, ZERO, or MAX) was quantified. This approach enables the reader to compare general trends within the cohort while still maintaining the focus on changes exhibited by individual students between the start and end of the complete laboratory sequence.

To observe how students' perceived knowledge *compared* to their actual knowledge, we looked at student responses to questions on the URSSA and LST that mapped to specific objective themes. Among the many objective topics for unit operations laboratory, the comparisons presented in this study focus on three key themes of **Safety, Data Analysis,** and **Familiarity with Equipment** because of their strong representation among both assessment instruments. The objective topics and the associated question description from these instruments are listed in Figure 1.

Value Rubric for Assessing Written Communication

To directly observe changes in students' technical writing communication skills, the initial report submissions (i.e. the reports that had been revised based on peer review feedback but had not yet received formal instructor feedback) were collected for the first project in both CHE Laboratory II and CHE Laboratory III. These artifacts were evaluated using the Written Communication VALUE Rubric.[23] Assessment was completed by four Rose-Hulman faculty members from departments other than chemical engineering. Raters worked in pairs to apply each criterion of the VALUE rubric to the submitted documents. Each pair was checked for interrater reliability during the first three report reviews and at every tenth report thereafter. Pairs were not permitted to advance further unless any conflicts that arose during these checks were resolved. Due to the nature of the artifact collection process

and the effort to maintain student anonymity, the data were reported by the evaluators in aggregate and represents data from all of the writing samples collected (Lab II N = 49, Lab III N = 39).

To see how students' technical communication skills changed compared with students' self-perceptions of their writing skills, results were compared with select URSSA instrument questions related to Written Communication, shown in Figure 2. The reports related to the first project of CHE Laboratory II provided the initial point of comparison, while the reports related to the first project of CHE Laboratory III provided the final point of comparison. Both of these reports are identical in style and are prepared individually. The final report of CHE Laboratory III could not be used for comparison as it is a long-form group report, which would make assessment of individual students difficult.

RESULTS

Average student scores for the different elements of the MUSIC Model for eMpowerment, Usefulness, Success, and Interest were compared pre-lab sequence and post-lab sequence. As shown in Figure 3, the students' average self-assessment scores were relatively high (above a 4 on a 6-point Likert scale) for all four categories before and after the laboratory sequence with slight movement in the average. Two sample Z-test results showed the average scores pre and post did not change in a statistically significant way (α = 0.05). Average values remain consistent between pre- and post-laboratory responses; however, the effect of averaging the scores may obfuscate clear changes in specific student attitudes.

The distribution of student rankings for each category was also investigated. The results are shown in Figure 4. By eliminating the average and breaking the data into number of scores, a clearer picture of the changes in students' attitudes can be measured. We observe in all four categories the mode rating for each category was a 5. In all categories other than Success, there were increases in some of the lower ratings (1, 2, 3), while there are no low ratings for Success at the end
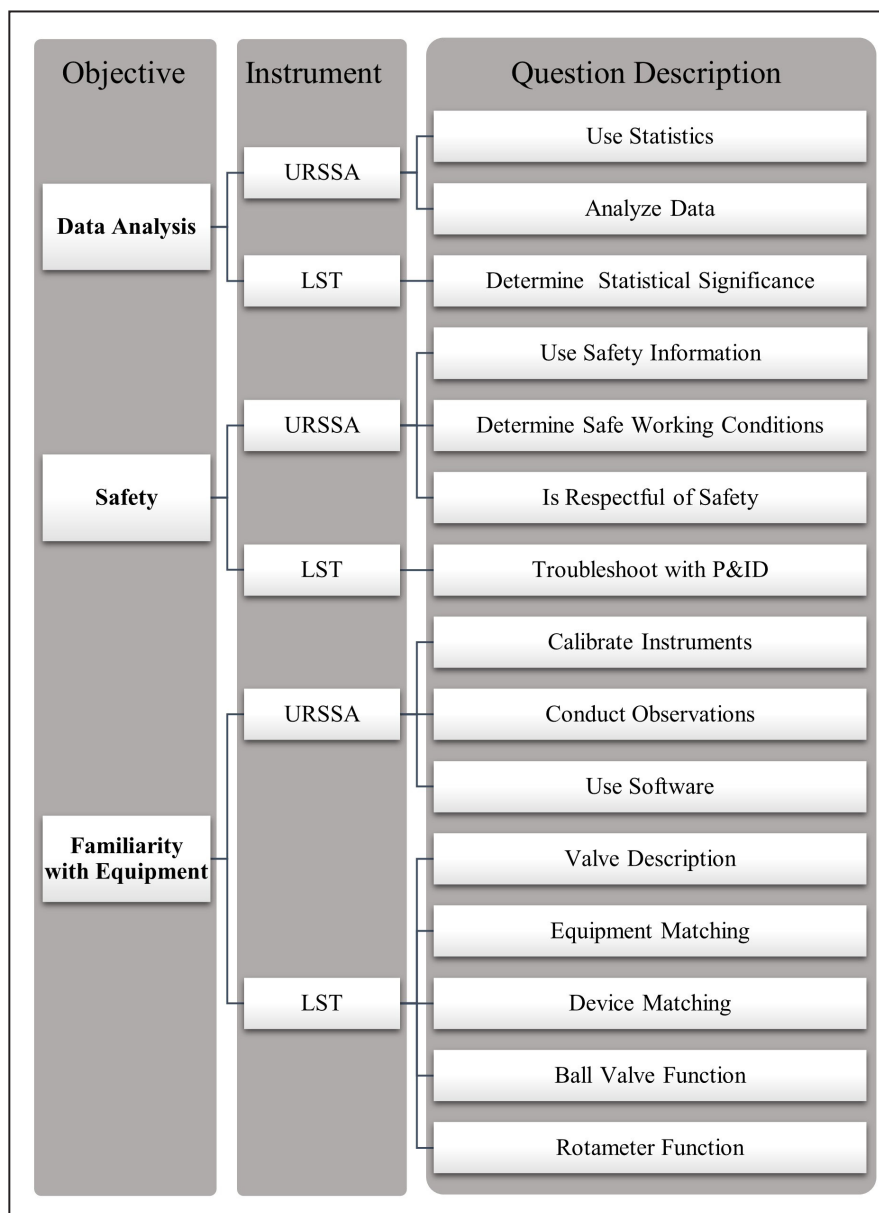

Figure 1. *Summary of objective topics and the associated assessment instruments and question descriptions. Where URSSA captures self-assessed knowledge; LST captures direct knowledge and abilities.*

of the course. This detail is not captured by simple averages shown in Figure 3. Unlike the other three categories, the Interest questions had more movement to lower scores, with 22% reporting a low rating compared to only 5% in the pre-sequence. Despite the increase in the number of students that scored their Interest as 1, 2, and 3, there were at least 75% of students scoring in the high range (4, 5, 6) post-sequence. Interest, in particular, showed an increase to the extreme ratings, with an increase of 6% in 1-2 ratings and an increase of 5% in 5-6 ratings. Interestingly, more than 13% gave ratings
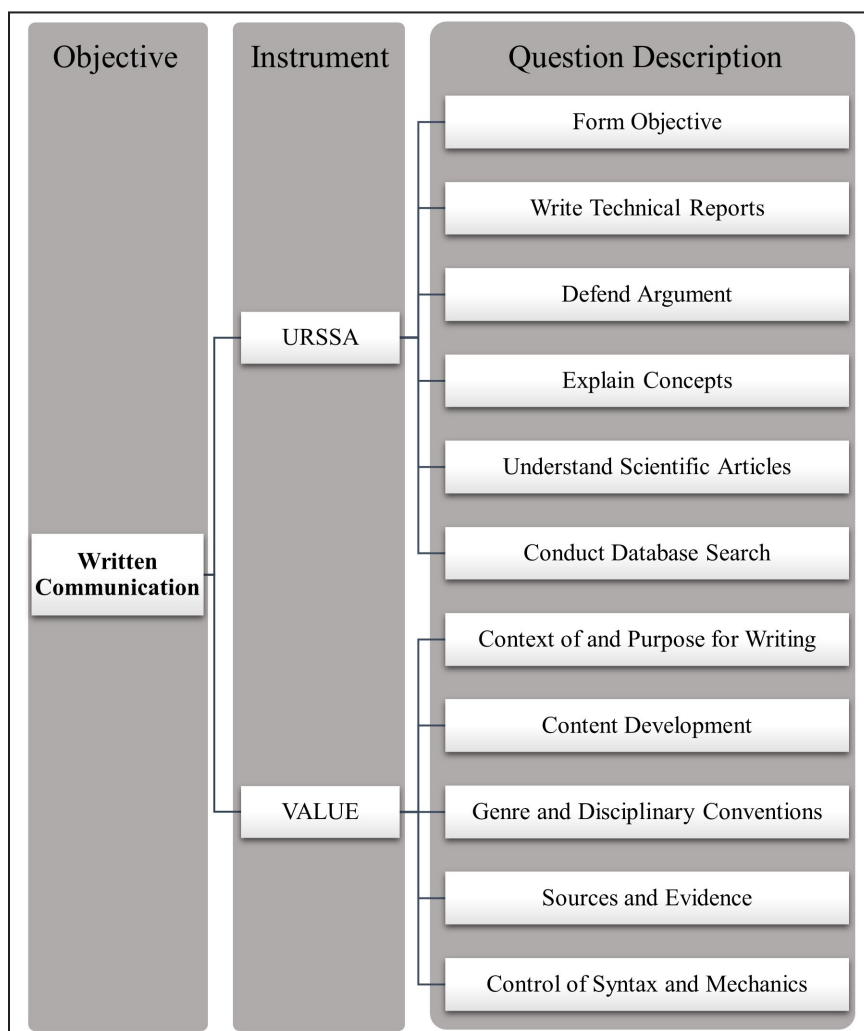
Figure 2. Summary of objective topics and the associated assessment instruments and question descriptions using self-assessment (URSSA) and direct assessment (VALUE) for written communication competencies.
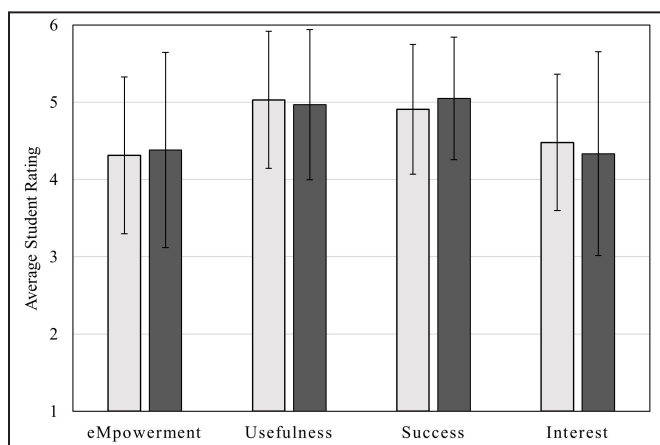

Figure 3. Comparison of averaged self-assessment data using the MUSIC Model before (light gray) and after the laboratory sequence (dark gray).

of 6 post-sequence for all four categories, and ratings of 6 only decreased in the category for Usefulness from pre (38%) to post (26%). Although there appear to be qualitative shifts in the distributions, Kolmogorov-Smirnov tests revealed that changes in distributions pre and post for each component were not significant ($\alpha = 0.05$). We acknowledge these statistics may be underpowered because of limited sample size. In conclusion, we observed no significant shifts to lower scores in any component, which we anticipated in our initial hypothesis.

By administering both self-assessments (URSSA) and direct assessments (LST), changes in student attitude and ability can be tracked, and student ability and self-perception of their ability can be directly compared. The distributions of student movement from select questions from the URSSA instrument and LST that mapped to the objectives of **Safety, Data Analysis**, and **Familiarity with Equipment** are shown in Figure 5.

Narrowing the focus, three themes were analyzed and compared between perceived and direct assessments. To visualize the results, the percentage of students was plotted versus the *movement*. Although Figure 5 shows that many of the themes had agreement between students' perceived movement and the direct assessment, there were instances where the two tools were not reconciled, which is consistent with previous studies.[10]

The theme of **Safety** yielded different trends depending on the question. Self-assessment questions regarding ability to perform a safety-related task returned positive movement among at least 50% of students. Interestingly, the question about respect for safety resulted in over half the students rating it the highest score (a 6 out of 6 on a Likert scale) before and after the sequence. This outcome could result from the emphasis on incorporating safety in the engineering curriculum where students gain a healthy respect for safety early in their studies.[24, 25] It also reinforces the important difference between attitude and ability. The direct assessment of safety in the LST showed the students' ability to safely operate equipment based on a P&ID improved over the course sequence, which agreed with their self-assessment. Similarly, the theme of **Familiarity with Equipment** resulted in agreement between perceived and direct assessment with positive movement for most questions. However, students
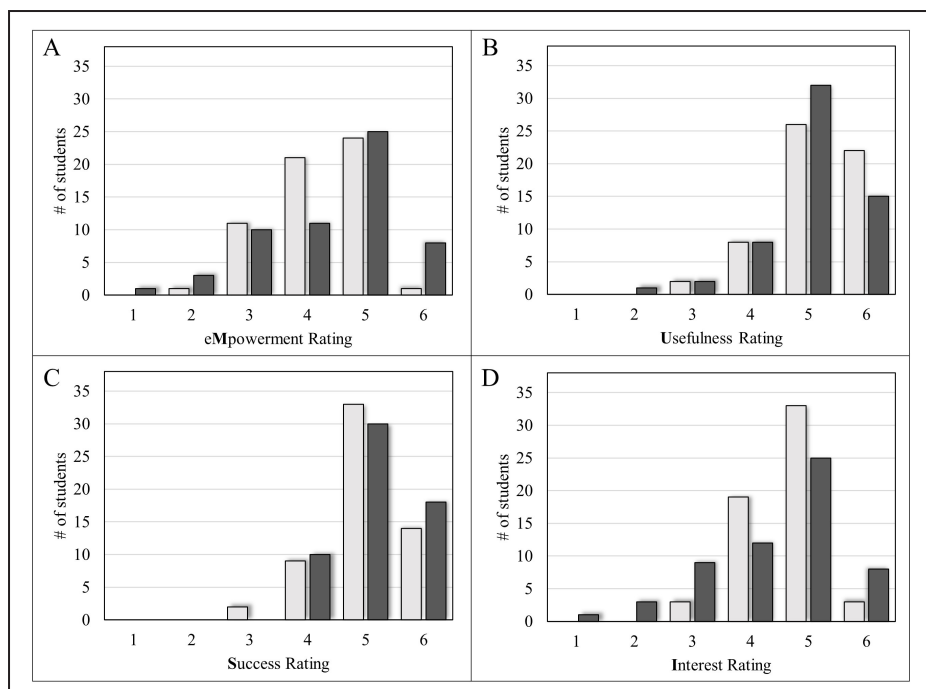
Figure 4. *Comparison of distributed self-assessment data using the MUSIC Model before and after the laboratory sequence. The lighter bars represent pre-laboratory scores, and the darker bars represent scores after the laboratory sequence was completed. Part A is eMpowerment ratings, B is Usefulness ratings, C is Success ratings, and D is Interest ratings. (N=58)*

still have difficulty appropriately reading a rotameter, as about half of the students showed negative or zero movement on this question.

Lastly, results from the **Written Communication** VALUE Rubric were compared with students' responses from the URSSA instrument related to this theme. Like the themes above, the students perceived a positive movement in all URSSA questions shown in Figure 6 regarding technical report writing skills. However, when assessed using the VALUE Rubric, the distribution of rubric scores was very similar across the two artifact collections, with the majority of all students receiving one of the two highest scores (3-milestone, 4-capstone). It is an acknowledged limitation that the study design lacked the ability to track individual student movement, as this may have revealed meaningful trends for individual students.

DISCUSSION AND CONCLUSIONS

In this study we explored how students subjectively feel about their learning in the laboratory compared to how much they actually learn and retain in a year-long CHE laboratory course. The common theme across the analysis was that, in general, student performance improved via direct assess-

ment, student self-perceptions of knowledge and abilities increased, and students registered positive attitudes about the course sequence throughout the assessment period. Incorporating multiple dimensions of assessment allowed for the formation of a more complete picture of laboratory outcomes informed by the distribution of student opinions and comparisons to directly assessed skills. Tracking student movement using these tools provides a better picture of the range of student experience.

The data show that the mode did not shift in the Interest category from the MUSIC Model; however, a higher number of students ranked their interest with low scores (1, 2, or 3) at the end of the sequence. This observation fit with anecdotal data regarding the laboratory sequence, where some students reported feeling that the three-course laboratory sequence seems repetitive or lost novelty over time. An interesting take away from this approach was that these themes did not necessarily match up with the most vocal student comments or complaints in course evaluations or via other feedback mechanisms. In general, there was a positive response, indicating most students find the course is useful, report that it captures their interest, and believe that they can be successful while feeling empowered in the course sequence. This aligns with our expectations, as we make a concerted effort to eliminate repetition by rotating projects, instructors, and teams while giving students repeated opportunities to hone their skills. However, by the end of the final course, there are additional scores of 1-3 where students are feeling less empowered, less interested, and feel the course is less useful. This change may be due to repetition, which can be a positive for honing skills, but not for keeping interest. Some additional influences could be from job offers and graduate school acceptances being received, causing shifting motivations, as well as coming to the realization that their "niche" in chemical engineering may not involve lab or the projects to which they were assigned. There may also be some commonality between the students who score lower post-sequence; an ongoing departmental effort explores why students are losing interest across the sequence to make informed improvements to our laboratory sequence. Incorporating these tools helps begin to clarify the noise from unmotivated or disinterested students who can get undeserving attention. Using these multiple assessment
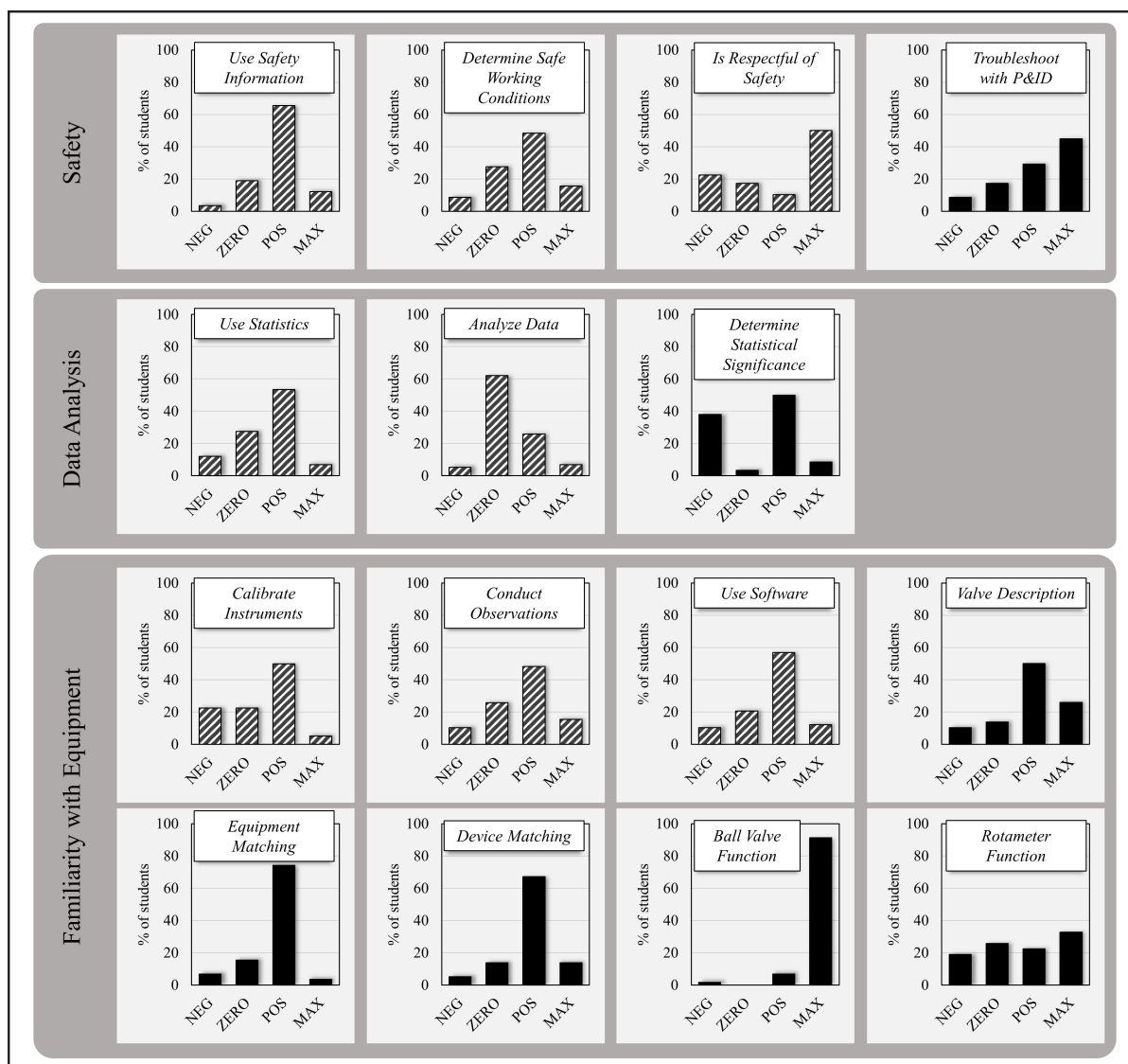
Figure 5. *Stripes represent movement in student perception from the self-assessment instrument (URSSA) and solids represent movement in student knowledge from the direct observation instrument (LST), N=58.*

methods allows educators to propose meaningful, data-driven changes to the laboratory curriculum, rather than attempting to make changes based on student feedback alone.

Although the overall themes were positive trajectories for most concepts in the URSSA and LST, there were notable discrepancies between what the students rated and what the direct assessment results showed. The differences may be due to the specific projects that the students performed in the laboratory. Although students complete five projects by the end of the sequence, there are some projects that specifically address some of the concepts that are assessed in the tools, while other projects do not. This gives some students more of an opportunity to explore and receive substantive feedback from faculty when writing their reports. For example, students

who work on the fluid flow experiment can conduct multiple repeat trials at the same set points. This type of experiment lends itself to more emphasis on statistical significance, and it is possible that the students on this type of project would score higher on the LST because they had direct experience with this concept. If a student is on the distillation project, it is possible that because the number of data points gathered is much smaller, their experience with statistical analysis will be much less extensive for this project. Similarly, there is some disagreement for the theme of familiarity with equipment. Although most projects have some form of valving that matched student perception and direct assessment, not all projects use a rotameter. Given the power of experiential learning,[26] this may be why some students improve on this
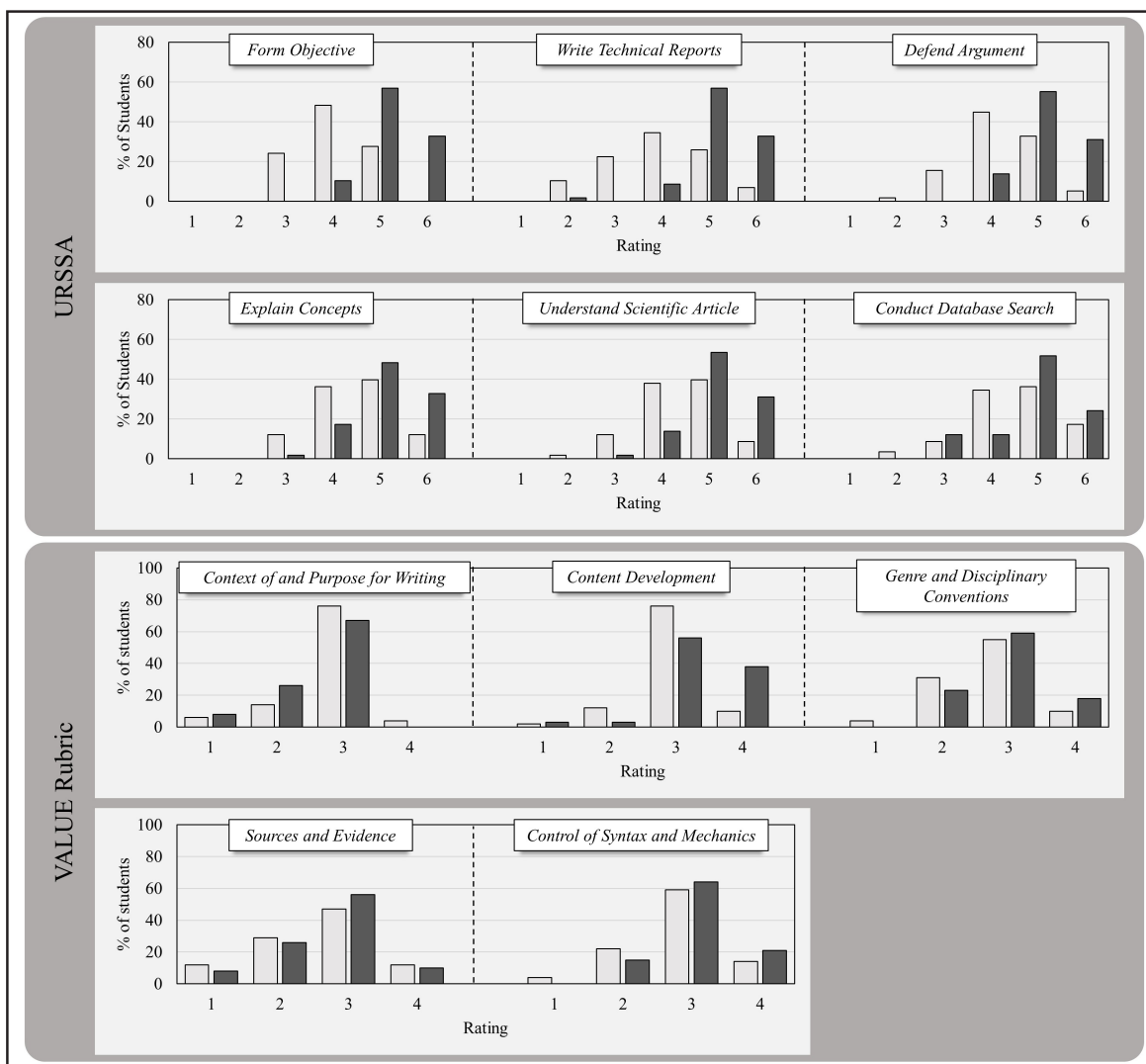
Figure 6. *Comparison of first (light gray) and last (dark gray) data collected from student self-assessment data obtained from select URSSA Instrument questions and direct assessment of artifacts using the Written Communication VALUE Rubric.*

question, while others did not. However, the knowledge gained from this assessment helped us identify this area as a potential knowledge gap within the laboratory sequence, allowing for it to be addressed in the future.

Throughout the multi-dimensional approach, the simultaneous deployment of the assessment tools gave a unique picture of the ability to assess both students' attitudes as well as measurable gains in knowledge and abilities. By using this overarching multi-dimensional framework for assessment, instructors have the power to uncover underlying gaps in student experience that may go unnoticed if the evaluation relies on just one form of assessment. Although the results presented assessed the initial use of the multi-dimensional approach for the unit operations laboratory sequence to gauge how students

developed throughout the courses, the implications of the tools used are farther reaching. Using the results of this study as a benchmark, many options for future use of these tools exist. While the multi-dimensional approach was demonstrated here, there is still an opportunity to select specific tools or a combination thereof to gain understanding of student attitudes or directly assess student laboratory knowledge. These tools can be implemented in various laboratory courses without many changes depending on the information that is desired by the instructor to assess and make data-driven changes to their laboratory courses. We are planning ongoing refinements to the Laboratory Skills Test to ensure proper interpretation of questions, such as a think-aloud study with students, that will further enhance the tool.

HOW CAN I USE THESE TOOLS?

The framework for this study made use of multiple assessments that aligned with the aim to observe student attitudes, knowledge, and abilities in a laboratory course via self-assessment and direct assessment. Table 1 shows how each instrument was used in this study in alignment with those aims. To use these assessment tools, consult the directing links and sources below.

URSSA:https://www.colorado.edu/eer/research-areas/undergraduate-research/evaluation-tools-undergraduate-research-student-self Customization of the survey can be executed similarly to the approach that was taken in this paper. Customization allows the surveyor to use the format of the original survey but focus on areas of interest.

MUSIC Model of Motivation Inventory: www.themusic-model.com. Use for an activity or course is free. Contact Dr. Brett D. Jones (brettdjones@gmail.com) for permission to use in publications.

VALUE Rubric; Written Communication: www.aacu.org/value-rubrics. All VALUE rubrics are available at no cost.

To obtain the customized URSSA prompts and the Laboratory Skills Test, please contact Dr. Neumann at neumann@rose-hulman.edu.

REFERENCES

1. Feisel LD and Rosa AJ (2005) The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education*. 94(1): 121-130. DOI: 10.1002/j.2168-9830.2005.tb00833.x.
2. Vigeant MA, Silverstein DL, Dahm K, Ford LP, Cole J, and Landherr LJ (2018) How we teach: Unit operations laboratory. *Proceedings of the ASEE Annual Conference*. https://peer.asee.org/how-we-teach-unit-operations-laboratory
3. Wright D, Markoff E, Grove J, Al-Hammoud R, Milne A, and Pirnia M (2017) Surveys for simultaneous direct and indirect Assessment. *Proceedings of the Canadian Engineering Education Association (CEEA)*. DOI: 10.24908/pceea.v0i0.10236.
4. Poulin J and Matis S (2015) Social work competencies and multidimensional assessment. *Journal of Baccalaureate Social Work*. 20(1): 117-135. DOI: 10.18084/1084-7219.20.1.117.
5. Williams TR (2009) The reflective model of intercultural competency: A multidimensional, qualitative approach to study abroad assessment. *Frontiers: The Interdisciplinary Journal of Study Abroad*. 18(1): 289-306. DOI: 10.36366/frontiers.v18i1.267.
6. Warburton EC (2002) From talent identification to multidimensional assessment: Toward new models of evaluation in dance education. *Research in Dance Education*. 3(2): 103-121. DOI: 10.1080/1464789022000050480.
7. Ciorba CR and Smith NY (2009) Measurement of instrumental and vocal undergraduate performance juries using a multidimensional assessment rubric. *Journal of Research in Music Education*. 57(1): 5-15. DOI: 10.1177/0022429409333405.
8. Jones BD (2009) Motivating students to engage in learning: The MUSIC model of academic motivation. *International Journal of Teaching and Learning in Higher Education*. 21(2): 272-285. https://www.isetl.org/ijtlhe/pdf/IJTLHE774.pdf
9. Hunter A-B, Weston TJ, Laursen SL, and Thiry H (2009) URSSA: Evaluating student gains from undergraduate research in the sciences. *CUR Quarterly*. 29(3). http://www.cur.org/documents/Citation_viewer/?Id=1340
10. Boud D and Falchikov N (1989) Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*. 18(5): 529-549. DOI: 10.1007/BF00138746.
11. Jones BD (2015) Motivating Students by Design: Practical Strategies for Professors. CreateSpace, Charleston, SC.
12. Jones BD and Skaggs G (2016) Measuring students' motivation: Validity evidence for the MUSIC model of academic motivation inventory. *International Journal for the Scholarship of Teaching and Learning*. 10(1): n1. DOI: 10.20429/ijsotl.2016.100107.
13. Lee WC, Brozina C, Amelink CT, and Jones BD (2017) Motivating incoming engineering students with diverse backgrounds: Assessing a summer bridge program's impact on academic motivation. *Journal of Women and Minorities in Science and Engineering*. 23(2). DOI: 10.1615/JWomenMinorScienEng.2017017960.
14. Virguez L and Reid K (2016) Students' perceptions in a first year engineering classroom and their relationship with behavioral and cognitive engagement. *First Year Engineering Experience (FYEE)* Conference. https://fyee.asee.org/FYEE2016/papers/150.pdf
15. Pace AC, Ham A-JL, Poole TM, and Wahaib KL (2016) Validation of the MUSIC® model of academic motivation inventory for use with student pharmacists. *Currents in Pharmacy Teaching and Learning*. 8(5): 589-597. DOI: 10.1016/j.cptl.2016.06.001.
16. Jones BD and Carter D (2019) Relationships between students' course perceptions, engagement, and learning. *Social Psychology of Education*. 22(4): 819-839. DOI: 10.1007/s11218-019-09500-x.
17. Laursen S, Hunter A-B, Seymour E, Thiry H, and Melton G (2010) *Undergraduate Research in the Sciences: Engaging Students in Real Science*. John Wiley & Sons, Hoboken, NJ.
18. Weston TJ and Laursen SL (2015) The undergraduate research student self-assessment (URSSA): Validation for use in program evaluation. *CBE—Life Sciences Education*. 14(3): ar33. DOI: 10.1187/cbe.14-11-0206
19. McDevitt AL, Patel MV, and Ellison AM (2020) Lessons and recommendations from three decades as an NSF REU site: A call for systems-based assessment. *Ecology and Evolution*. 10(6): 2710-2738. DOI: 10.1002/ece3.6136.
20. (2010) Directions for preparing and administering *URSSA*. https://www.colorado.edu/eer/sites/default/files/attached-files/urssadirectionsfeb2010.pdf accessed August 2020.
21. Ribera T, Anastasio DD, Chenette H, and Neumann GT (2018) Work in progress: Developing a multi-dimensional method for student assessment in chemical engineering laboratory courses. *Proceedings of the ASEE Annual Conference*. https://peer.asee.org/30014
22. Koretsky MD, Falconer JL, Brooks BJ, Gilbuena DM, Silverstein DL, Smith C, and Miletic M (2014) The AIChE "Concept Warehouse": A web-based tool to promote concept-based instruction. *Advances in Engineering Education* 4(1): n1. https://advances.asee.org/wp-content/uploads/vol04/issue01/papers/AEE-13-Milo-cor2.pdf
23. Association of American Colleges and Universities (AAC&U) (2009) Written Communication VALUE Rubric. Retrieved from https://www.aacu.org/value/rubrics /written-communication accessed August 2020.
24. Lemkowitz SM (1992) A unique program for integrating health, safety, environment and social aspects into undergraduate chemical engineering education. *Plant/Operations Progress*. 11(3): 140-150. DOI: 10.1002/prsb.720110308.
25. Dembe AE (1996) The future of safety and health in engineering education. *Journal of Engineering Education*. 85(2): 163-168. DOI: 10.1002/j.2168-9830.1996.tb00226.x.
26. Kolb DA (1984) *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, Englewood Cliffs, NJ. ❏