



EvaluateUR: Helping Students to Acquire the Knowledge and Skills Needed for Success in the Workplace

JILL SINGER^{1*}, DANIEL WEILER², JOHN DRAEGER¹, JILL ZANDE³, SEAN FOX⁴, and BRIDGET ZIMMERMAN⁵

¹*SUNY Buffalo State University, Buffalo, NY 14222, USA*

²*Daniel Weiler Associates, Kensington, CA, 94707, USA*

³*MATE Inspiration for Innovation, Monterey Peninsula College, Monterey, CA, USA*

⁴*Science Education Resource Center, Carleton College, Northfield, MN 55057, USA*

⁵*Nautilus Evaluation Services, Rushville, NY 14544, USA*

**singerjk@buffalostate.edu*

Abstract: The EvaluateUR Method supports the assessment of undergraduate research programs in different educational settings, from independent undergraduate research (“EvaluateUR”) to classroom-based research (“EvaluateUR-CURE”) and robotics design competitions (“Evaluate Compete”). The method provides statistically reliable assessments of student growth in a wide variety of outcome categories identified as essential to success in the workplace. It differs from traditional approaches to assessing student outcomes because it is integrated directly into the research experience. A unique feature of the method is its emphasis on metacognition. Thus, it also serves as a learning tool for students, helping them to become more aware of their academic and professional strengths and weaknesses while supporting their efforts to identify strategies for expanding their knowledge and improving their metacognitive skills.

Keywords: EvaluateUR method, Assessment, Undergraduate Research, Metacognition, Workplace Skills

© 2023 under the terms of the J ATE Open Access Publishing Agreement

Introduction

Industries and businesses across the country have identified technical and professional competencies and skills they consider essential in today’s workplace [1,2,3]. These skills include communication, problem-solving, time management, and teamwork. Above all, industries and businesses want employees who can identify and think through problems and determine how to solve them. If students are to succeed as employees (or as entrepreneurs), they will have to master these and related skills, be aware of what they know and don’t know, and understand how best to overcome any weaknesses in their knowledge and skills.

Undergraduate research programs have proven their value over many years at colleges and universities nationwide [4-7]. These experiences provide students with many important insights and skills related to their academic interests and the process of systematic inquiry.

We developed a method [8,9,10] that provided reliable data on the value of undergraduate research across a wide range of desirable skills. Faculty representing both STEM and non-STEM disciplines identified a list of undergraduate research outcome categories of interest and contributed to developing the components used to define each outcome category. We were particularly interested not only in how well students did on research projects, but in student knowledge, skills, and abilities desired in the workplace. The method - known as EvaluateUR - was designed to gather statistically reliable empirical data on student outcomes and also to ensure student awareness of the range of skills that employers value, sharpen student insight into their strengths and weaknesses, and provide students with the self-reflective and analytic tools they will need to succeed. To accomplish these goals, the evaluation employs an assessment instrument that is completed by both the student researchers and their research mentors at different times during the students’ research projects. The instrument covers a wide range of student skills – including many “soft” skills valued by employers – and becomes the basis for student-mentor conversations to discuss reasons for their respective assessment scores, critically examine the degree of student insight into their academic strengths and weakness and consider potential



student strategies for leveraging strengths and overcoming weaknesses. EvaluateUR encourages students to be aware of what learning strategies they employ and why and use that awareness to make adjustments that help them learn more effectively. This cycle of self-awareness, adjustment, and renewed self-evaluation – widely known as **metacognition** – is an essential element in the EvaluateUR method. It is reinforced by a set of separate exercises that provide students with additional opportunities to develop their metacognitive skills. Acquiring the habit of metacognition is perhaps the single most important benefit that students take away from their EvaluateUR experience.

EvaluateUR was shown to be effective [9], and with funding from the NSF WIDER program was scaled up and pilot tested with undergraduate research programs at over 40 colleges and universities across the country [9,10]. Our evaluation found that EvaluateUR introduced students to a wide range of competencies and skills that are valuable in education and the workplace; measured student growth in mastering those competencies and skills; contributed to the development and enhancement of metacognitive skills; enabled mentors to focus their efforts on areas where students were weak; and helped students gain new insights into their academic strengths and weaknesses. EvaluateUR is now available by subscription and provides undergraduate research program directors with reliable evidence to document the benefits of their programs.

Since its introduction and with funding from the NSF ATE program, the method has seen two adaptations: EvaluateUR-CURE (E-CURE) supports students enrolled in course-based undergraduate research experiences (CUREs), and Evaluate-Compete (E-Compete) supports students participating in remotely operated vehicle (ROV) regional and international competitions created by the Marine Advanced Technology (MATE) Center and now organized and administered through MATE Inspiration for Innovation (MATE II). While E-Compete's initial design is intended for ROV competitions, it can be adapted to other student team efforts under the direction and mentoring of an advisor, including other robotics competitions and events such as the Community College Innovation Challenge. The need for variants stems primarily from the differences in student-advisor/mentor ratios, the duration of the research experiences, and the need to align outcomes to rubrics used by competition judges. Thus, while the three variants of the method support a broad range of research experiences, they are more alike than different. The underlying approach is the same and embeds metacognition, encouraging students to self-reflect on their strengths and weaknesses. Differences in implementation steps and the number of outcomes recognize differences in research settings but do not compromise the method's original intent or basic design.

Description of the EvaluateUR Method

All three variants of the EvaluateUR method include a set of outcome categories (Table 1), and each outcome category is defined by several components. Examples of outcome categories and defining components are included in Table 2, with a complete list of the outcome categories and components found at <https://serc.carleton.edu/evaluateur/methods/outcomes>. All three variants of the method share common key features. These include: 10 or 11 student outcome categories with options to add several additional outcomes; each outcome category is delineated by several components that measure specific outcome objectives; repeated assessments at the beginning, middle, and end of the research so that students' progress can be followed; independent faculty/mentor assessments and student self-assessments using identical instruments and a 5-point rubric based on how often the student has exhibited the behavior described by a particular component (1=Not Yet to 5=Always); and student-faculty/mentor conversations to improve students' metacognitive insights into their strengths and weaknesses. In addition to the outcome categories and components listed in Table 1, E-Compete includes a set of ROV-specific outcome categories and components that align with the scoring rubric used by judges during the competition, such as: vehicle design, buoyancy, and propulsion; control and electrical system; sensors, payload, and tools; safety; project management; and entrepreneurship.



Table 1. List of Outcome Categories for the Three EvaluateUR Variants

Outcome Categories	EvaluateUR	EvaluateUR-CURE	Evaluate-Compete
Communication	✓	✓	✓
Creativity	✓	✓	✓
Autonomy	✓	✓	✓
Ability to Deal with Obstacles	✓	✓	✓
Intellectual Development	✓	✓	✓
Critical Thinking and Problem Solving	✓	✓	✓
Practice and Process of Inquiry	✓	✓	
Nature of Disciplinary Knowledge	✓	✓	
Project Knowledge and Skills	✓	✓	✓
Teamwork / Collaboration		✓	✓
Ethical Conduct	✓	✓	

Table 2. Examples of Outcome Components for Three Outcome Categories

Outcome Categories	Outcome Components for each Variant		
	EvaluateUR	EvaluateUR-CURE	Evaluate-Compete
Ability to Deal with Obstacles	Is not discouraged by setbacks or unforeseen events and perseveres when encountering challenges.		
	Shows flexibility and a willingness to take risks and try again.		
	Trouble-shoots problems and searches for ways to do things more effectively.	Demonstrates ability to quickly improvise & implement a solution to fix a design or equipment problem	
Critical Thinking & Problem Solving	Looks for the root causes of problems and develops or recognizes the most appropriate corrective actions.		
	Recognizes flaws, assumptions, and missing elements in arguments	This outcome component is not used in Evaluate-Compete	
	This outcome component is not used in EvaluateUR or EvaluateUR-CURE	Demonstrates the ability to evaluate alternative designs and/or operational solutions	
Project Knowledge & Skills	Displays knowledge of key facts and concepts.	Displays an understanding of the engineering and scientific principles and practices relevant to vehicle design and operation	



2.1 Steps in the Implementation of the EvaluateUR method

To illustrate the implementation sequence for the EvaluateUR method, Figure 1 shows the four stages of EvaluateUR and the steps that take place during each stage. E-CURE and E-Compete both follow the same sequence, as explained below.

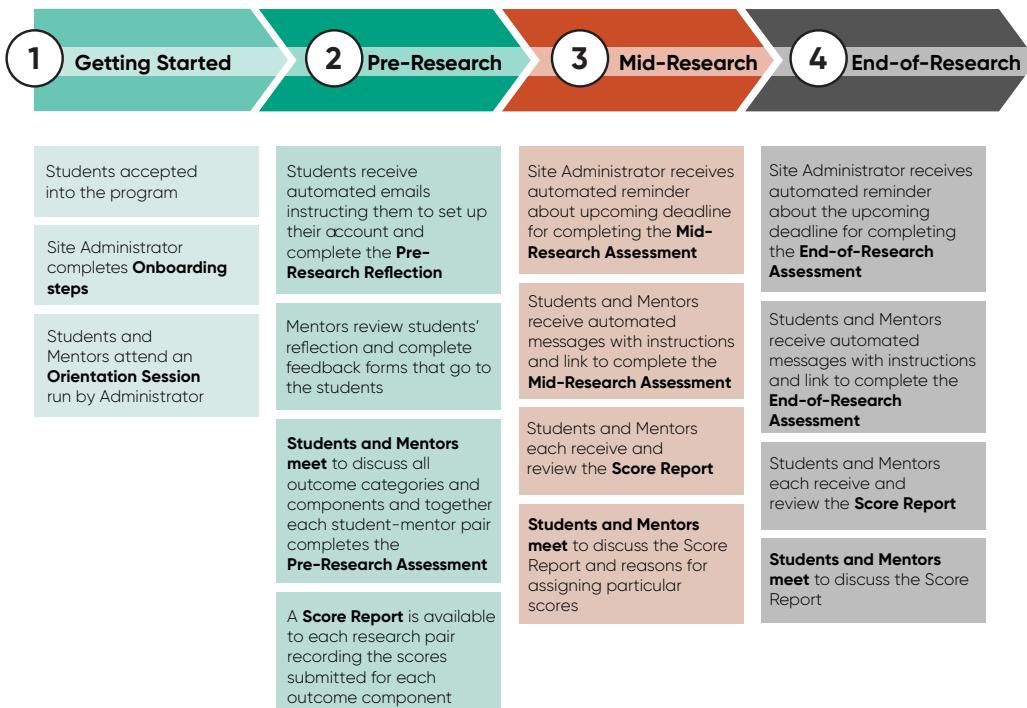


Figure 1. Four stages of EvaluateUR

1.) Getting Started

Site administrators (usually program directors or coordinators) learn how to use the EvaluateUR method by completing the **onboarding** steps that illustrate and explain how to configure their program dashboard (e.g., setting assessment completion dates, adding student-mentor pairs, adding optional outcomes or questions). Once it is set up, the dashboard shows all the student-mentor pairs and the sequence of steps to be completed by each pair. Clicking inside the box for any student-mentor pair expands the box and lists what actions happen at each step of the process. As steps are completed, they change color from gray to orange to green, making it easy to see what steps are completed (green), what action is required to complete the current step (orange), and steps that have not yet been started (gray). This helps the administrator track the progress of each student-mentor pair. The last activity is running an **orientation session** to help students and mentors understand the purposes and advantages of EvaluateUR. For E-CURE and E-Compete, the CURE instructor or ROV team advisor completes the onboarding process to learn how to set up the dashboard and introduce the method to their students. For all variants, selecting the 'activate' option on the dashboard results in an automated message alerting students to begin the Pre-Research steps.

2.) Pre-Research

Students record their ideas about the research process by answering a set of open-ended questions called a **pre-research reflection**. This is intended to provide mentors with more information about their students, and the student responses can highlight any concerns the students might have about the research they are about to begin. After mentors review the responses, the students and mentors complete the **pre-research assessment** for the list of student outcomes (see Table 1). The outcome scores allow students and mentors the opportunity to exchange ideas about the importance of the various outcomes. To facilitate these conversations, score reports are automatically generated and provide a side-by-side comparison of the scores assigned by the students and mentors for each outcome component. This makes it easy to identify outcome components with score differences of 2 or more (based on the 5-point rubric).



Depending on the variant, the pre-research assessment is done when the student-mentor pair meets to discuss each outcome category and its components (EvaluateUR), or is completed only by the student (E-CURE and E-Compete). The different approaches to completing this pre-research assessment reflect the diverse research environments for the three variants. For example, the one-to-one student-mentor ratio in EvaluateUR makes it feasible for the student-mentor pair to complete the pre-research assessment together. In contrast, the larger number of students enrolled in a CURE or a member of an ROV team makes this impractical. In E-CURE, students are reminded that the assessment scores are not considered in their grades but provide an independent picture of student growth on the outcome measures. For E-CURE and E-Compete, the course instructor or team advisor, respectively, have access to a data summary of all the students' scores for each outcome category.

3.) Mid-Research

Mid-research assessments are completed independently by students and mentors about halfway through the research. Following this, automated messages are sent with a link to a score report that provides the scores for pre- and mid-research assessments. Each student-mentor pair then meets to discuss the reasons for assigning particular scores. This provides an opportunity for students to consider how they might leverage their strengths and adopt strategies to help them tackle areas of weakness as they continue their research. For E-CURE and E-Compete, these meetings might involve a class-wide discussion, meetings with research groups, and/or meetings with individual students. There also is an option in E-CURE and E-Compete to select a sub-set of outcome categories to be used on the mid-research assessments. This sub-set can be selected at the same time the CURE instructor or ROV team advisor sets up their dashboards and can be modified at any time before releasing the mid-research assessment. This option accounts for the higher number of students CURE instructors/ROV team advisors are mentoring and help them concentrate on observing fewer outcomes as they interact with all the students.

4.) End-of-Research

This stage in the process is very similar to what happens at the mid-point. For EvaluateUR, students and mentors complete the **end-of-research assessments** and schedule the final conversations. For E-CURE and E-Compete, students answer a set of open-ended questions following the completion of the end-of-research assessments. These questions help students reflect on their experiences and provide CURE instructors and ROV team advisors with another way to assess student learning. This can help them consider how they might modify their pedagogical strategies. In addition, E-Compete includes an optional de-brief exercise that ROV team advisors can use soon after the students return from the ROV competition.

2.2 Impacts of the EvaluateUR Method

Based on statistical analyses of data and responses to student and mentor surveys, a number of findings about the impacts of the method are clear. A key result is that the method's structure supports meaningful dialogues between students and mentors/instructors/ROV team advisors. For a great majority of the students using the three variants of the method, most outcome scores improved over time. A summary report from the final year of the NSF WIDER project collected data on 799 student-mentor pairs representing STEM and non-STEM disciplines. For all outcome components, there was a statistically significant increase in student growth in assessment scores given by students and mentors when analyzed by a paired sample t test ($\alpha = .05$). An analysis of effect size (Cohen's d) showed medium and large magnitudes of effect for almost all components. The medium and large effect sizes suggest that the significance was not due to chance or large sample size but to actual impacts on student outcomes. Student and mentor survey responses confirmed that the repeated conversations contributed to developing and enhancing student metacognitive skills. This is characterized by learners becoming aware of what learning strategies they are pursuing and why, and then using that awareness to make intentional adjustments to those strategies to learn more effectively. The conversations also helped most students confirm their plans for continuing their education at the graduate level or seeking employment in their discipline. Research mentors found it easier to identify the academic strengths and weaknesses of their students, enabling the mentors to better focus their guidance. The mentors could also more easily identify areas where students might be over- or under-estimating their abilities and were able to help students gain new insights into their academic strengths and weaknesses and the relative efficacy of their learning strategies. Research mentors also reported that using EvaluateUR contributed to changing their attitudes about what students are capable of doing, leading them to rethink their pedagogical practices.



Findings from a two-semester-long CURE taken by engineering technology students confirm that E-CURE outcome categories and components correspond to ABET ETAC performance indicators [4]. Integrating E-CURE into this design course has provided regular and structured feedback to the students, including serving as an early warning system should individual students or a team of students fall behind in meeting project deadlines. E-CURE also supports the generation of individual and whole-class assessment scores that align with ABET performance indicators. According to E-CURE instructors teaching other STEM courses, this variant promotes a positive learning environment and helps students become more resilient in overcoming research obstacles. E-CURE also provides data that help instructors revise their CUREs to improve student learning.

Student survey feedback from all three variants indicates that the key project design features – introduction to outcomes important to employers, independent student/mentor assessments, follow-up conversations, and metacognitive exercises – have introduced students to new ideas and have helped them to think strategically about skills they need to sharpen or acquire. In addition, faculty survey responses have indicated that EvaluateUR and its variants have encouraged them to make changes to their courses to incorporate more intentional metacognitive growth activities and discussions about the skills employers value.

3. Metacognition and the EvaluateUR Method

The EvaluateUR method encourages students to be aware of what learning strategies they employ and why [11]. Then, they can use that awareness to make adjustments that enable them to learn more effectively [12, 13, 14]. As noted above, this cycle of self-awareness, adjustment, and renewed self-evaluation, or metacognition, is perhaps the most important benefits students can take away from their EvaluateUR experience.

To further support the development of metacognition, a collection of 12 exercises has been developed. These exercises provide additional practice that supports the metacognitive benefits of the method. Table 3 provides a brief description of each exercise with downloadable full versions found at <https://serc.carleton.edu/evaluateur/methods/outcomes.html>.



Table 3. Metacognition Exercises

Title of Exercise	Brief Description
Learning From Past Projects	Students reflect on how they've navigated past projects and assignments. The goal is to help students learn from those experiences and develop the independence necessary for a research project.
Building Project Management Skills	Students reflect on how they are currently navigating their research project.
Thinking About How to Ask Good Questions	Students reflect on how they formulate questions as well as how they generate answers. The aim is to prompt students to think about disciplinary modes of thinking and what constitutes appropriate evidence.
Thinking About How to Ask Good Research Questions	Students reflect on how they formulate questions central to their research and what counts as adequate evidence. The aim is to prompt students to connect disciplinary modes of thinking with research projects.
Building Resilience	Help students reflect on how they can overcome obstacles. It asks them to think back to a prior experience and draw out lessons that might help them succeed in their research projects.
Building Research Resilience	Students reflect on how they're coping with setbacks related to the research process.
Reading with a Purpose	This activity is designed to be used in conjunction with a reading assignment. The aim is to prompt students to read more intentionally and draw out lessons that might help them succeed in their research projects.
Reading for Research	Students reflect on how they're doing the reading related to their research project.
Thinking About Self-Assessment Process	Encourages students to reflect on how they completed the self-assessment to consider whether it was fair and accurate.
Better Together: Teamwork and Collaboration	Asks students to reflect on how teams can function effectively and get collaborations back on track when they run into trouble. The aim is to prompt students to think about how to have a good research team experience.
Thinking About How You Communicate	Asks students to reflect on how to effectively express their work to a disciplinary audience. The aim is to encourage students to develop clear, concise, and organized modes of communication.
Thinking About How You Communicate Across Audiences	Asks students to reflect on how they communicate with others about their research project. The aim is to prompt students to express ideas in a clear and concise manner using discipline-specific language.

Each exercise is short and follows a format that includes an introduction to students about the purpose of the exercise. This is followed by several questions intended to help the student reflect on strategies they have used in the past that might be useful in assessing a particular situation and identifying how they might tackle that situation. Metacognition exercises are not intended to be graded. Instead, they are designed to help the students think about strategies in their 'toolbox' and how to use them effectively. A user guide for the exercises is available on the website and is intended to provide adopters of the EvaluateUR method with an overview of metacognition and how to help their students build and apply their metacognitive skills. Because E-Compete is likely used for students participating on an ROV or equivalent team, a metacognition card game has been developed to replace the more traditional exercises developed for E-CURE and used for EvaluateUR. The card game has 30 cards divided into three major categories: People, Problem-Solving, and Persistence. Each card poses a situation that is intended to help students think about how they approach a particular situation or express their thoughts. The card game can be used as an 'icebreaker' for ROV team members to get to know each other better, and to introduce metacognition into break time when the team members are not focused on tasks related to designing, building, and testing their vehicle.



Conclusions

The EvaluateUR method differs from other undergraduate research assessments because it provides a framework for student-mentor conversations that are aimed at helping students understand their strengths and areas where improvements are needed. By including measures of outcomes valued in the technical workplace the method assesses a diverse range of student knowledge and skills that go beyond those of immediate interest to specific research projects. The three variants of the method serve specific research settings (e.g., independent research, course-based research, and robotic competitions). Each variant is integrated directly into the research experience, thereby providing assessments of student outcomes that serve as measures of student success and as learning tools for the student. A primary benefit of the method is that it encourages students to become more aware of what learning strategies they employ to analyze and solve problems. It also strengthens their ability to recognize situations where they need to learn new skills and/or seek assistance from others. For mentors, adopting and implementing the EvaluateUR method has contributed to a greater awareness of the value of structured feedback. In some cases, it has resulted in changes to their pedagogical strategies.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grants No. 1347681, 1347727, 1836033, 1932929, and 1932940. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Disclosures. The authors declare no conflicts of interest.

References

- [1] N. H. Hensel, *Undergraduate Research at Community Colleges: Equity, Discovery, and Innovation*, (Stylus Publishing 2021), 229 p.
- [2] H. Huvard, R. M. Talbot, H. Mason, A. N. Thompson, M. Ferrara, B. Wee, “Science identity and metacognitive development in undergraduate mentor-teachers,” *Int. J. STEM Educ.* 7(1) 1-17 (2020).
- [3] National Association of Colleges and Employers (NACE), <https://www.naceweb.org/career-readiness/competencies/>.
- [4] I. Grinberg, J. Singer, “ETAC ABET and EvaluateUR-CURE: Findings from combining two assessment approaches as indicators of student learning outcomes,” *J. Eng. Technol.* 8-22 (2021).
- [5] H. A. Mieg, E. A. Ambos, A. Brew, D. M. Galli, and J. Lehmann, *The Cambridge Handbook of Undergraduate Research*, (Cambridge University Press 2022) pp. 729.
- [6] J. Gentile, K. Brenner, and A. Stephens, *Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities*, (National Academy Press 2017) pp. 290.
- [7] H. J. Schmitz, K. Havholm, “Undergraduate Research and Alumni: Perspectives on learning gains and post-graduation benefits,” *CUR Quarterly* 35(3) 15-22 (2015).
- [8] J. Singer, D. Weiler, “A Longitudinal Student Outcomes Evaluation of the Buffalo State College Summer Undergraduate Research Program,” *CUR Quarterly* 29(3) 20-25 (2009).
- [9] J. Singer, B. Zimmerman, “Evaluating a Summer Undergraduate Research Program: Measuring Student Outcomes and Program Impact,” *CUR Quarterly* 32(3) 40-47 (2012).
- [10] J. Singer, D. Weiler, B. Zimmerman, S. Fox, E. Ambos, “Assessment in undergraduate research,” in H. A. Mieg, E. Ambos, A. Brew, D. M. Galli and J. Lehmann (Eds.), *The Cambridge Handbook of Undergraduate Research*, (Cambridge University Press 2022) pp. 158-171.
- [11] L. Scharff, J. Draeger, D. Verpoorten, M. Devlin, L. S. Dvorakova, J. M. Lodge, and S. Smith, “Exploring Metacognition as Support for Learning Transfer,” *Teaching & Learning Inquiry* 5(1) 1-14 (2017).



- [12] K. K. Karukstis, T. E. Elgren, *Developing & Sustaining a Research-Supportive Curriculum: A Compendium of Successful Practices*, (Council on Undergraduate Research, 2007) pp. 598.
- [13] G. Schraw, K. J. Crippen, and K. Hartley, “Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning,” *Research in Science Education* 36 111-139 (2006).
- [14] K. D. Tanner, “Promoting student metacognition,” *CBE—Life Sciences Education* 11(2) 113-120 (2012).