

General Education Diversity Assessment Committee (GEDAC)
Final Report of 2017-2018 Pilot Project Activities
September 5, 2018

Report Prepared by:
Cynthia Reyes, PhD
Katharine Shepherd, EdD
Alex Yin, PhD

GEDAC Final Report of 2017-2018 Pilot Project Activities

Background

The purpose of the General Education Diversity Assessment Committee (GEDAC) is to assess the extent to which students are achieving the identified competencies and outcomes identified in relation to the UVM Diversity Curriculum Requirement initially approved by the Faculty Senate in 2006 and implemented in 2007. A revised set of competencies was approved by the Faculty Senate on May 14, 2015 and forms the basis of the current assessment effort. Initial assess Final Report of 2017-2018 Pilot Project Activities assessment activities conducted during the 2016-2017 academic year included gathering data on enrollment patterns in D1 and D2 courses, conducting a pilot survey of faculty, and assisting with the design and analysis of a student focus group process conducted through CTL during the spring of 2017. As 2017 came to a close, our focus shifted to the charge to conduct a direct assessment of student outcomes as evidenced by their written work in a sample of D1 courses. We were invited to seek funding from the Provost's Office, and in August, 2017, the Provost approved a grant to support the design, implementation, and evaluation of activities associated with this assessment effort. Funds were used to support a doctoral student (Arby Ghemari) who helped to organize meetings, take notes, and collect feedback on the rubric used for the project. Grant funds were also used to support faculty members who participated in the Rating Day and subsequent data analysis occurring over the summer of 2018. 2018 GEDAC members include Cynthia Reyes and Katie Shepherd (co-chairs, CESS), and Matt Kolan (RSNER), Abby McGowan (CAS), Cathy Paris (Faculty Senate), and Marie Veal-Fagnant (RSNER). Additional faculty/instructors who participated in some or all of the activities associated with the design and implementation of the GEDAC Pilot Project include Mercedes Avila (CNHS), Heather Edy (CESS), John Gennari (CAS), Ginny Hu (CAS), Winnie Looby (CESS), Michael McDonald (RNESR), and Zach Ispa-Landa (RSNER). Importantly, Alex Yin from the Office of Institutional Research (OIR) supported all aspects of the GEDAC's work.

The remaining sections of this report summarize the activities of the GEDAC and associated D1 faculty who participated in the Pilot Project during the 2017-2018 academic year, and describe the outcomes of our mixed methods approach to assessing student outcomes in relation to one of the 14 diversity curriculum competencies.

Methods

Design and Implementation of the Pilot Project

The design phase of the GEDAC Pilot Project began in the Spring of 2017 when the GEDAC invited Libby Miles to talk about her experiences assessing outcomes associated with the FWIL General Education Requirement. Margaret Burke (RSNER) also provided information regarding the curriculum alignment and assessment process undertaken by

faculty in RSNER. Based upon this information, as well as familiarity with rubrics available through the American Association of Colleges and Universities (AAC&U), the GEDAC agreed that the most feasible approach for the pilot would be to identify one to three competencies, select D1 instructors whose courses addressed those competencies, develop a rubric, and apply the rubric to a sample of student work from Fall 2018 D1 classes. The committee identified two primary goals for the GEDAC project: first, to use a rubric to evaluate the quality of student work, for the purpose of determining the degree to which UVM students are achieving learning outcomes associated with the diversity curriculum competencies; and second, to evaluate the efficacy of the pilot process itself. These goals were critical, as the GEDAC was fully aware of the complexity of the task before us, and the need to identify an assessment approach that could be replicated over time in all D1 and D2 courses, across all 14 competencies.

During the Fall of 2017, the GEDAC held five 90 minute meetings devoted to clarifying the parameters of the pilot project, identifying the competencies to be addressed, identifying potential rubrics for consideration, identifying and recruiting potential faculty participants, and (ultimately) revising an AAC&U rubric on Intercultural Competence to be shared with and revised by pilot project faculty in advance of the targeted May 30 Rating Day. GEDAC co-chairs and Alex Yin met five additional times (60 minutes each) in the fall semester to document the decisions being made, work with the doctoral student on tasks that needed to be completed, revise materials as needed, and plan agendas for the monthly meetings. A few key outcomes emerged from these meetings. First, the GEDAC determined that it would focus its assessment on only one competency area, Awareness (See Table 1 for a list of the 2015 competencies), which consists of four separate competencies. This decision was made in recognition of the difficulty of developing an assessment system capable of evaluating multiple competencies. Additionally, the committee made an assumption that a majority of D1 courses would, regardless of their specific content area, need to address issues of awareness as they pertained to key concepts around diversity. Second, the GEDAC agreed that the AAC&U rubric on Intercultural Competence would provide the best starting point for developing a rubric that could be used with ease across courses earmarked as D1 courses. The committee spent two meetings revising the original AAC&U rubric prior to sharing it with the additional faculty members who made further revisions prior to the Rating Day. Finally, the GEDAC decided to send an open invitation to all D1 faculty who taught courses in the Fall of 2018 to determine who might be willing to participate in a series of Spring meetings and (optionally) the Rating Day in May. Nine faculty (hereafter referred to as GEDAC faculty associates) responded to the invitation and joined GEDAC monthly meetings beginning in January, 2018.

Monthly 90-minute meetings held during the spring semester included discussions of the purpose and nature of the pilot project, as well as discussions and activities designed to identify additional revisions to the rubric. For example, during two meetings, the GEDAC and faculty associates read and rated samples of student work from a D1 class taught by one member of the GEDAC. Ensuing discussions about areas of agreement and disagreement in rating served as an initial norming process and resulted in further

revisions to the rubric. Table 1 depicts the 14 UVM competencies, and Figure 1 shows the original rubric in relation to the four Awareness competencies.

Table 1: The University of Vermont Diversity Competencies and Intended Outcomes

Competencies
Awareness
1. Develop an awareness of the diversity of individuals, cultures, and communities within the U.S and globally.
2. Develop an awareness of one’s identities, attitudes, beliefs, values, and assumptions.
3. Develop an awareness of how one’s identities, attitudes, beliefs, values, and assumptions influence how one interacts with or views those who are similar to or different from oneself.
4. Develop an awareness of the influence and impact of diversity-related concepts such as intersectionality (i.e., the intersection of multiple dimensions on diverse cultural identities), power and privilege, and so forth at the individual, group, and systems levels.
Knowledge
5. Cultivate knowledge and an understanding of the histories, traditions, values, contemporary issues, experiences, demographics, and practices linked to diverse identities and groups within the U.S. and globally.
6. Cultivate an understanding of foundational concepts such as diversity, culture, equity, equality, social justice, power/privilege, etc.
7. Cultivate an understanding of the origins and systemic nature of prejudice, injustice, discrimination, and oppression directed toward people of diverse backgrounds.
8. Develop critical thinking skills to enable one to analyze information and evaluate arguments from diverse viewpoints and multiple perspectives.
9. Develop the written and oral communication skills necessary to engage in intellectual discourse about diversity-related topics.
10. Develop interpersonal skills that support respectful, meaningful, and effective interactions with those from diverse backgrounds, including understanding the intent and impact of one’s actions.
Integration and Application
11. Develop the ability to recognize and name dynamics (e.g., identity-related, cultural, power) at the individual, group, and systems levels that are present when exploring issues related to diversity.
12. Engage skills to problem-solve and to develop action plans that address diversity-related issues and situations.
13. Develop the ability to effectively intervene and respond to acts of bias or intolerance, as well as to resolve conflicts that occur across differences.
14. Develop the ability to effectively facilitate the learning and development of others around diversity.

Figure 1: Rubric with Four Awareness Competencies

**General Education Diversity Assessment Rubric
Draft: December 15, 2017**

To be used in evaluating student work gathered in D1 classes

	1	3	3	4
DEGREES OF Awareness	Emerging	Reaching	Addressing	Exceeding
Develop an awareness of the diversity of individuals, cultures, and communities within the U.S. & globally.	Evidence of <i>surface</i> understanding of the diversity of individuals, cultures, and communities within the U.S. & globally.	Evidence of <i>partial</i> understanding of the diversity of individuals, cultures, and communities within the U.S. & globally.	Evidence of <i>adequate</i> understanding of the diversity of individuals, cultures, and communities within the U.S. & globally.	Evidence of <i>sophisticated</i> understanding of the diversity of individuals, cultures, and communities within the U.S. & globally.
Develop an awareness of one's identities, attitudes, beliefs, values, and assumptions.	Evidence of <i>surface</i> awareness of one's identities, attitudes, beliefs, values, and assumptions.	Evidence of <i>partial</i> awareness of one's identities, attitudes, beliefs, values, and assumptions.	Evidence of <i>adequate</i> awareness of one's identities, attitudes, beliefs, values, and assumptions.	Evidence of <i>sophisticated</i> understanding of one's identities, attitudes, beliefs, values, and assumptions.
Develop an awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	Evidence of <i>surface</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	Evidence of <i>partial</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	Evidence of <i>adequate</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	Evidence of <i>sophisticated</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.
Develop an awareness of the influence and impact of diversity-related concepts such as intersectionality, power and privilege, and so forth at the individual group, and systems levels.	Evidence of <i>surface</i> awareness of the influence and impact of diversity-related concepts such as intersectionality, power and privilege, and so forth at the individual group, and systems levels.	Evidence of <i>partial</i> awareness of the influence and impact of diversity-related concepts such as intersectionality, power and privilege, and so forth at the individual group, and systems levels.	Evidence of <i>adequate</i> awareness of the influence and impact of diversity-related concepts such as intersectionality, power and privilege, and so forth at the individual group, and systems levels.	Evidence of <i>sophisticated</i> awareness of the influence and impact of diversity-related concepts such as intersectionality, power and privilege, and so forth at the individual group, and systems levels.

Figure 2: Rubric for Reviewed Student Work in Relation to One Awareness Competency

	0	1	2	3
Awareness Competencies (UVM D1 competencies)	No Evidence	Superficial	Deepening	Sophisticated
Develop an awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	No evidence	Evidence of <i>superficial</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself.	Evidence of <i>deepening</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself. Evidence of awareness of the forces and experiences which have shaped one's personal identity, attitudes, beliefs, values and assumptions.	Evidence of <i>sophisticated</i> awareness of how one's identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar to or different from oneself within the U.S. & globally. Evidence of awareness of how one's identities, attitudes, beliefs, values and assumptions intersect with power and privilege

In subsequent meetings, norming activities and discussions highlighted the diversity of courses and range of competencies that instructors may have or may not have emphasized in their courses; as such, the committee decided to move away from evaluating student work in relation to the four Awareness competencies and decided instead to use the rubric in relation to just the third Awareness competency. This competency asserts that students will “Develop and awareness of how one’s identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar or different from oneself.” As seen in Figure 2, the final version of the rubric also included a score of “0” to indicate that student work showed “No Evidence” of the selected Awareness Competency. Each participating faculty member was charged with identifying a key written assignment within their class and submitting de-identified student work that responded to that assignment to the GEDAC.

Rating Day

In preparation for the May Rating Day, GEDAC co-chairs, along with Alex Yin, met to plan the agenda and prepare Rating Day materials, including the revised rubric (Figure 2). Participation in the Rating Day had been described as optional for faculty associates, and in the end, four faculty associates (including two co-instructors of a larger D1 course) submitted student work from three courses to be used in the Rating Day. One other faculty associate participated in Rating Day but did not submit student work. Alex Yin spent approximately a half-day identifying the representative sample of student work to be evaluated on Rating Day (see below for a full description of the sample), de-identifying the work and assigning codes, and making copies for the Rating Day. Student codes were linked to UVM ID numbers, allowing subsequent analyses of ratings based on students’ school year, gender, and race/ethnicity.

The Rating Day was held from 9:30 – 4:30 on May 30, with participants including the five faculty associates, three GEDAC members, and Jennifer Dickinson (Assessment Faculty Fellow, CTL). Following an overview of the day, J Dickinson led the group in a 90-minute norming activity, which included application of the rubric to two pieces of student work that had been submitted but were not included in the Rating Day sample. The ensuing discussion resulted in determination of an acceptable level of inter-rater agreement as well as agreement that the “0” rating would be used when student work did not provide any evidence of awareness as indicated on the rubric, and/or when their work simply did not address awareness as defined by the rubric. For example, a well-constructed and well-written student assignment could still receive a “0” if it included objective information on a diversity-related topic but did not explore issues of awareness as defined in the targeted competency. Following norming, the instructors who had provided student work gave a brief overview of their D1 class and the instructions given to students in advance of their work. This allowed the group to have a clearer understanding of the overarching goals of each class and the nature of the student assignment. As will be elaborated upon in the Results section, participants noted at the outset that the three assignments varied quite a bit, both in their format and in the degree to which they directly asked students to talk about their awareness of key elements of diversity.

The process used by participants to rate student work included the following:

Two sub-groups were identified (4 in one group and three in the other) and given copies of the 69 papers identified for the sample. As noted, the sample included student work from three distinct D1 classes. This process ensured that every paper was read by at least two Rating Day participants (one in each of the two sub-groups). Raters used the final version of the rubric to score student work. Each piece of student work was assigned a code, and raters indicated their rating score (0 – 3) on a master list.

At the conclusion of the Rating Day, participants wrote individual responses to four process-oriented questions asking them to reflect on: (a) one or two things they learned from the rating process; (b) one or two things they learned from reading student work; (c) utility/efficacy of the rubric; and (d) ways in which the Rating Day activities might inform future approaches to assessing student outcomes in D1 and D2 courses.

Written responses were collected for analysis purposes. Additionally, the group engaged in a 45-minute discussion of the Rating Day process, including overall impressions, positive and challenging aspects of the process, and implication for larger assessment efforts. Notes of this conversation were recorded by the co-chairs and were analyzed along with individual written reflections.

Sample

Using data available through OIR, a representative study sample (N= 69) was derived from the total population of students (N= 163) for whom we obtained work in three distinct D1 courses. Within the sample, 62% of students were female and 38% were male, corresponding to the gender breakdown in the population (61% female and 39% male). Table 2 shows the breakdown of the study sample by race and ethnicity.

Table 2: Race/Ethnicity Proportions of Population and Sample

Race/Ethnicity	Population	Sample
Total	163	69
Asian	1.8%	4.4%
Black/African American	2.5%	2.9%
Hispanic	6.7%	8.7%
Two or more	1.8%	44%
White	77.3%	60.9%
International	8.0%	15.9%
Unknown	1.8%	2.9%

In both cases, the majority of the students were White, though the sample reflected a higher percentage of international students than the population (16% as compared to 8%). With respect to the class level¹ of students enrolled in the three courses, both the sample

¹ Class level is determined by the number of student credit hours earned at the beginning of the semester

and population reflected higher percentages of first-year students (41% and 50%, respectively) than any other category (Table 3).

Table 3: Class Level Proportions of Population and Sample

Race/Ethnicity	Population	Sample
Total	163	69
First-year	49.7%	40.6%
Sophomore	27.6%	31.9%
Junior	10.4%	17.4%
Senior	11.7%	10.1%
Other	.6%	0%

Data Analysis

Analysis of Rating Day data occurred on three levels. First, Alex Yin analyzed the scoring sheets and calculated frequencies to determine trends and patterns across key demographic characteristics of students in the sample; namely, gender, race/ethnicity, and class level (i.e., first-year to Seniors). Additionally, he analyzed inter-rater reliability in terms of scores assigned to the 69 papers by two readers each. These results were reviewed and discussed by Alex Yin and the GEDAC co-chairs, and constitute the Pilot Project's *assessment of student outcomes* in association with Awareness Competency #3. Next, the GEDAC co-chairs analyzed qualitative data as provided by GEDAC members and faculty associates in their written reflections, identifying major themes and patterns of responses. These results provided the basis for the *assessment of the pilot process*, including implications for future assessment processes. Finally, GEDAC members and faculty associates were offered the opportunity to read this report and to comment on the results section. This provided an opportunity for a member check, particularly with respect to our conclusions regarding the efficacy of the pilot process.

Inter-rater reliability

An analysis of inter-rater reliability indicated that, in spite of norming activities conducted prior to and on the day of the rating activity, significant differences did exist between raters. Of the 69 papers reviewed by two raters, no difference was indicated for 30 papers; however, a difference of 1 point (on a 0 – 3 point scale) was present for 36 papers, and a difference of 2 points was present for 3 papers. Possible reasons for these differences in inter-rater reliability were discussed during the de-briefing session on Rating Day and will be described in the Results and Limitations sections.

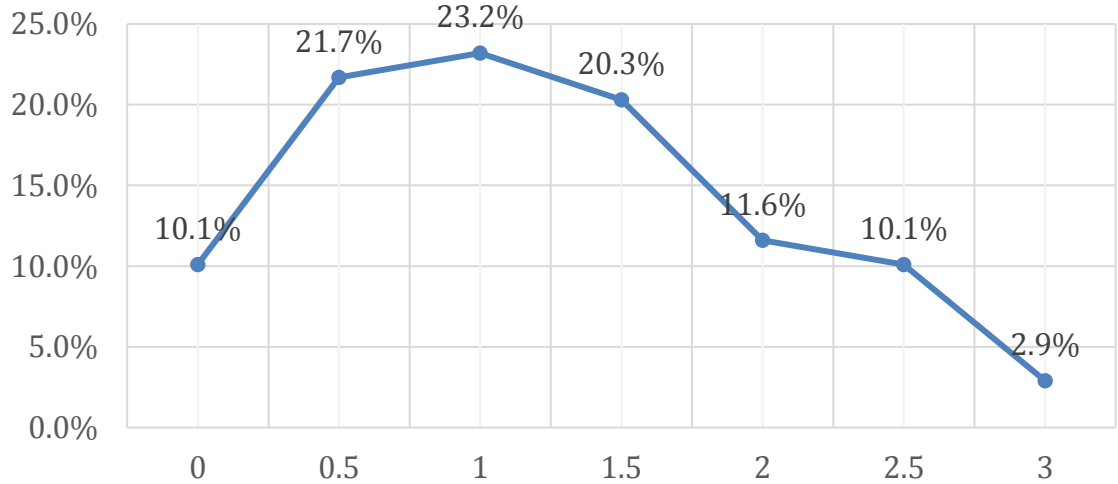
Results

Assessment of Student Outcomes

Figure 4 presents results of the frequency distribution calculated for the average score for each of the 69 pieces of student work. A holistic score of 1.22 (SD= .70) represents the

average across all student work, with 10.1% of all student work rated as a 0 and only 2.9% of student work receiving the highest score of 3.

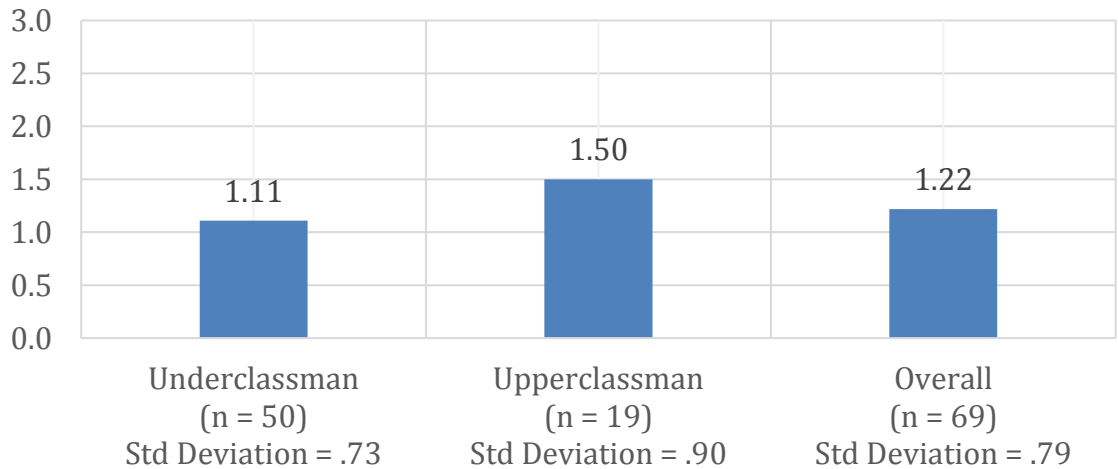
Figure 4: Frequency Distribution of Average Awareness Score (n=69, mean = 1.22, standard deviation = .70)



Analysis of student’s awareness of diversity by class level (see Figure 3) shows some variation between classes. Mean scores for first year students were 1.13, with mean scores for sophomores dropping a bit to 1.08, and mean scores for juniors and seniors rising to 1.46 and 1.57, respectively.

Figure 3: Awareness of Diversity by Class Level*

* No significant difference between underclassman and upperclassman students (F= 1.16 p-value = .333)

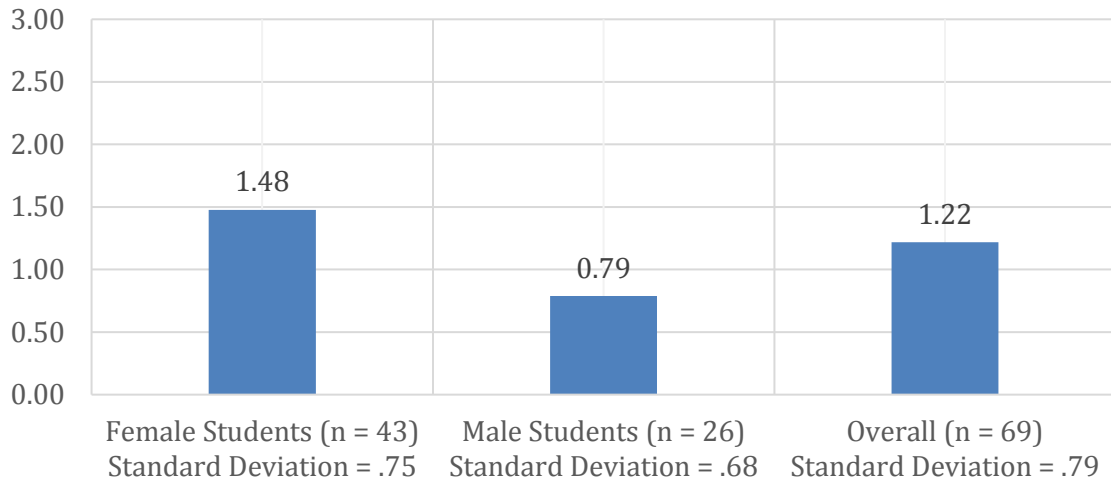


* No significant difference between underclassman and upperclassman students (F= 3.45 p-value = .068)

Although this general trend suggests that students’ awareness of diversity may be increasing over time as a function of their class level at UVM, no significant trend was

found when scores were compared for underclassmen (first-year students and sophomores) with upperclassmen (juniors and seniors). Significant differences were reported, however, among female and male students (see Figure 4), with average ratings for female students at 1.48, as compared to .79 for male students.

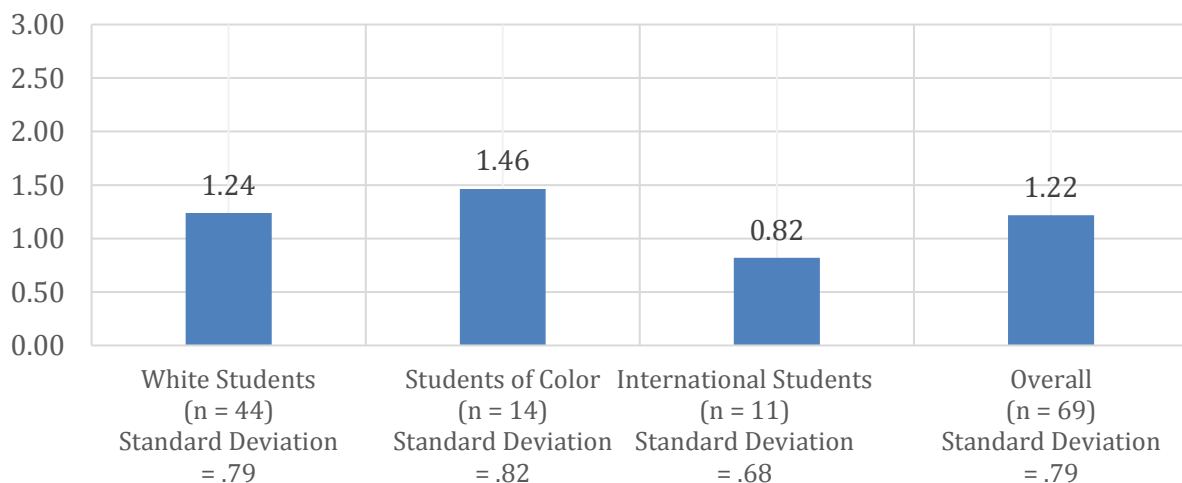
Figure 4: Awareness of Diversity by Sex*



*Significant difference between female and male students ($F= 14.67$ $p\text{-value} = <.001$)

No significant differences were reported for White versus non-White students, with average scores for White students at 1.24 ($N= 44$) and average scores for non-White students at 1.46 ($N=14$). Lower average scores (.82) were reported for international students ($N=11$). See Figure 5 for a summary of these scores.

Figure 5: Awareness of Diversity by Race/Ethnicity*



* No significant difference exists between groups ($F= 2.16$ $p\text{-value} = .123$)

Finally, a regression analysis was conducted to look for differences across groups (i.e., sex, student level, students of color, and international students), with significant differences observed for only female students (at the .001 level). The variables explained

28.2% of the variance in the awareness of diversity metric. Table 4 summarizes the results of the regression analysis.

Table 4: Regression Results of Awareness and Sex, Student Level, and Race/Ethnicity*

	Unstandardized B	Coefficients Std. Error	t	Significance
Constant	.727	.156	4.66	<.001
Sex				
Male = Ref. Group				
Female	.703	.178	3.95	<.001
Class Level				
Underclassman = Ref. Group				
Upperclassman	.404	.188	2.15	.036
Race/Ethnicity				
White = Ref. Group				
Students of Color	.048	.220	.22	.827
International Students	-.429	.235	-1.82	.073

*R-square = .282

Assessment of Pilot Process

Assessment of the pilot process occurred through an analysis of the post-rating reflections that participants (i.e. GEDAC members and faculty associates) submitted and the debriefing notes recorded by the GEDAC co-Chairs during Rating Day. Overall, it appeared that participants were highly enthusiastic of the process and felt that this particular retreat was valuable for advancing D1 and D2 initiatives at UVM. One participant noted the ways in which the discussion leading up to and occurring during Rating day helped to “see everyone’s perspectives on how they defined certain terms and expectations,” related to their D1 courses. Another participant felt the norming protocol was just as valuable as the actual rating activity, and a third observed how students do take the D1 work seriously and that their “cultural competency development extends far beyond a single artifact or class” during their time at UVM.

Along with their generally positive comments on the process, participants raised questions and acknowledged a number of challenges that can be summarized into four important learning points: 1) adoption and application of the 14 diversity competencies; 2) the challenges of defining awareness; 3) the relationship of assignments to assessment outcomes; and 4) considerations regarding use of the rubric.

Adoption and application of the 14 Diversity Competencies. During the spring semester planning sessions, as well as during the Rating Day, participants talked frequently about general challenges associated with the 14 Diversity Competencies. These included concerns that not all faculty teaching D1 and D2 courses had begun using the competencies adopted in 2015. Additionally, participants acknowledged their sense that there may be a lack of consistency among D1 and D2 faculty in terms of their willingness to regularly develop and use assessment tools. These concerns have been

noted in earlier initiatives of the GEDAC (i.e., the 2017 student focus groups and the GEDAC pilot faculty survey of D1 courses). Although further exploration of this issue is warranted, it appears unclear how D1 faculty across the University identify, refer to, and assess the D1 competencies in their courses. The GEDAC is encouraged by points made by Pablo Bose in sections 4 and 5 of the DCRC summary report for 2017-2018 to regularly assess D1 and D2 courses. This plan will greatly enhance understanding about how D1 faculty are incorporating the competencies in their classes.

The challenge to define “awareness.” Along with general concerns about adoption of the Diversity Competencies, participants’ written reflections and final discussion identified specific challenges related to defining the Awareness competencies in ways that will result in objective and consistent assessments of student outcomes. Among these challenges, participants expressed differences of opinion regarding definitions of awareness and the degree to which the university holds common expectations regarding the learning outcomes expected for all students. The participants wondered, for example, whether or not there is a universal expectation that participation in the diversity curriculum should result in all students achieving a “Sophisticated” level of awareness, or whether it would be acceptable to learn that students demonstrated a range of levels of awareness even after participation in D1 and D2 courses. Although the GEDAC selected the Awareness competencies on the premise that awareness would be addressed in most, if not all, D1 classes, our discussions of the competency area indicated that it may in fact be challenging to gain a common understanding of the meaning of awareness and how progress towards increasingly sophisticated levels of awareness can be evidenced in a single course. Calculations of interrater reliability supported this concern, with raters achieving “no difference” in scoring on only 30 of 69 papers, and a difference of 1 point (on the three point scale) for 36 papers. One participant suggested that the range of experiences among students taking D1 classes might contribute to the challenge of defining awareness consistently. As an example, this participant spoke about the contrast between White students who often talked about their “transformation” in a D1 class (e.g., describing a dramatic shift in their thinking based on a reading or class activity) and students of color who sometimes felt extremely uncomfortable because their own experiences and initial levels of awareness were at such different and often deeper levels than those of some of their White peers. When this occurred, it seemed that students of color could perceive claims of “transformation” by White students as superficial or contrived. Another participant commented that some students of color experienced their own levels of awareness as more sophisticated than their instructors’ awareness or competency, particularly if instructors were seen as relying too heavily on students of color to share their personal experiences in the classroom. The point of these examples is not to suggest that a particular experience is more authentic than another; rather, they emphasize the challenge of understanding what it means for students to grow in their levels of awareness given that they enter the university with varying degrees of emotional, cognitive, and experiential understanding of race and racism in the U.S. The participants’ consensus was that what matters most in a D1 course may be the degree to which students’ thinking changes as a result of completion of the course and/or the degree to which the course creates an environment and a level of cognitive dissonance about race and racism that fosters shifts in students’ awareness and thinking about these

issues. The degree to which this conclusion is shared by faculty outside of this group is a matter in need of further exploration, but it does point to a need for continued conversations about the purpose of the diversity curriculum requirement and university-wide expectations regarding student learning outcomes.

Relationship of assignments to assessment outcomes. The third learning point relates to the degree to which certain types of assignments are suited to the assessment of a complex construct such as awareness and to the particular competency addressed in this pilot (i.e., assessment of the degree to which “one’s identities, attitudes, beliefs, values and assumptions influence how one interacts with or views those who are similar or different from oneself”). Through the norming process it was apparent that a well-written paper that included conceptual knowledge of race and racism might still fail to demonstrate the complex interactions occurring between identity, attitudes, beliefs, values and interactions with others. In other words, a student could summarize critical historical events related to racial issues, but might fail to discuss how their personal identity and positionality mediated these issues. The participants engaged in a lengthy discussion of how the nature of particular course assignments appeared to influence the degree to which students’ work reflected varying levels of awareness. One of the assignments used in this process included an explicit focus on the degree to which students’ awareness changed from the beginning to the end of the course, one asked students to reflect on their understanding of diversity but did not include a direct prompt regarding their awareness in relation to their identity(ies), and the third assignment required students to discuss themes common to several class readings but did not require any sort of personal commentary or reflection. At the outset, the GEDAC’s intent was to create a rubric that could be applied to a range of assignments within all D1 courses, but it appeared that variation across the assignments may be problematic and in this case, may have contributed to variations among raters during the scoring process and questions about the degree to which the rubric could reliably measure student learning outcomes. A student might have in fact develop a greater sense of awareness through a D1 course experience, but if the assignment did not require specific attention to that issue, it may not have been reflected in their work or the resulting rating.

Rubric considerations. The fourth learning point expanded on the third and related to the rubric itself. The participants agreed that the rubric was well constructed, facilitated their understanding about what students learned in their various courses, and created an excellent vehicle for discussing the nuances of the Diversity competencies and related student outcomes. At the same time, they were not entirely convinced that the tool could produce reliable and valid results across courses. As one participant noted, “given the diversity of student work it can be like comparing apples and oranges.” Here too, participants recognized that an assessment of levels of awareness would be more valid if the assignment was geared toward actually gauging student awareness. The point seems obvious, but again raised the challenge of developing an assessment tool that would be both simple and flexible enough to be used across courses and instructors involved in the D1 and D2 curriculum.

In summary, the four main learning points emerging from participants' written reflections and discussions included 1) general challenges and questions related to the adoption and use of the 2015 Diversity Competencies; 2) the challenges of objectively assessing the concept of awareness, especially in light of students' varied backgrounds and experiences with race and racism; 3) the challenge to develop assignments that can accurately evaluate students' levels of awareness, particularly given the range of content addressed in D1 courses; and 3) the challenge to develop a rubric that would be simple and flexible enough to be readily adopted by instructors in all D1 courses, yet also sophisticated enough to reliably evaluate student learning outcomes across a wide variety of assessments. Participants also expressed familiar questions and concerns related to the vulnerability of students of color in large D1 discussion classes, the vulnerability of pre-tenure faculty of color who teach them, and a call for more leadership emphasis on motivating all faculty, and not just D1 and D2 faculty, to engage in professional development related to diversity courses. Each of these comments reflected the challenge of assessing student learning outcomes in the absence of an understanding of the classroom climate in which learning is occurring. GEDAC members and faculty associates had multiple conversations about the fact that the task of the committee was to design and evaluate the effectiveness of a process for evaluating student learning outcomes on a particular set of assignments, rather than to evaluate the pedagogical and environmental factors that served as the context for learning. At the same time, participants acknowledged that it was difficult to separate issues of pedagogy and classroom climate/management from discussions of student learning. They confirmed the need to continue to conduct university-wide conversations about classroom climate because of the strong potential of pedagogy and classroom management to influence student learning. These questions and concerns are integrated into the recommendations that participants gave for furthering D1 work included in the discussion section below.

Discussion

The following section highlights important implications and recommendations from this pilot study. At the outset of the process, the GEDAC's goal was to pilot a process to evaluate student learning outcomes and to report those outcomes along with feedback on the process itself. Results obtained on both dimensions need to be interpreted with caution, especially given the small sample size (69 pieces of student work from three D1 classes, evaluated by a total of nine D1 faculty and GEDAC members), the self-selection of faculty members into the process, and the fact that the group achieved an acceptable rate of interrater reliability for only 30 of 69 papers. At the same time, there is much to learn from the process.

With respect to the direct assessment of student learning outcomes, a few trends were noted, included a general trend for higher levels of awareness for students in their 3rd and 4th years at UVM as compared to students in their 1st and 2nd years; a significant difference between ratings for female and male students, with female students scoring higher than males; and lower scores for international students. No significant differences were found for White students as compared to non-White students. These findings must be considered in the context of the pilot study, but they raise some important questions

about the degree to which gender, international status, and year of study may be related to students' levels of awareness of diversity as measured by the targeted competency. The points raised in the post-rating discussion surfaced some challenges that need to be considered as the university continues to pursue a feasible approach to evaluating student outcomes on Diversity Competencies, including the pros and cons of requiring instructors to develop particular types of assignments in the evaluation of specific competencies, and the challenge to develop rubrics that can be consistently applied across a wide variety of instructors and courses.

The discussions occurring among GEDAC members and faculty participants about the pilot process itself proved to be an invaluable part of the project. During these discussions, faculty members were motivated to unpack and critique assumptions they may have had about a number of things, including their assessments, their expectations about assignments, their perceptions of the meaning of "awareness," and their concepts of what evidence would need to be demonstrated in student work in order to determine varying degrees of awareness regarding race and racism. As they shared these assumptions, participants became more noticeably attuned to the challenges of assessing student learning. They realized that beyond the differences they had in their beliefs about what and how to assess was the importance of cultivating a classroom climate that fostered student learning at whatever point students were on the continuum of learning. Additionally, they engaged in deep discussions about the purpose of the diversity curriculum requirement, surfacing questions about the degree to which the curriculum is designed to ensure all students achieve the highest level of proficiency articulated in the competencies, or whether variations in student learning were to be expected. This question seems particularly important when considering the complexity of the diversity curriculum requirement and the types of competencies it addresses.

Evaluation of the process also points to the time intensive labor of this type of assessment activity. On the one hand, the conversations emerging throughout the year served as a form of inquiry and professional development for all involved. GEDAC members and faculty participants alike were generally highly engaged and appreciative of the opportunity to explore questions related to assessment with peers from around the university. On the other hand, the process involved a great investment of time on the part of the GEDAC committee, the Office of Institutional Research, and faculty associates, resulting in concerns about the degree to which this approach to assessment can or should be replicated across all instructors, courses and diversity competencies.

Recommendations

Results of the pilot process lead us to several recommendations and thoughts about opportunities for continued exploration of approaches to assessing the diversity curriculum requirement.

1. Continue to examine processes for evaluating student learning outcomes related to the diversity curriculum requirement, taking into account the resources currently available to support the assessment of student learning outcomes in D1 and D2

courses. This report highlights the many benefits associated with the pilot project; at the same time, scaling this approach to address learning outcomes in all D1 and D2 courses and across all 14 competencies appears difficult, if not impractical, even if it continues to be done with a small sample. The GEDAC acknowledges that without the support provided this year through the Provost's Office, as well as ongoing support through the Office of Institutional Research, it would have been substantially more difficult to implement the pilot assessment project as conceptualized. Going forward, we recommend a careful review of the processes and resources needed to assess student learning outcomes across all competencies and D1 and D2 courses. The results of the review, along with the results of this pilot process, should then be used to guide the design of future assessment activities. Ideally, the review will be done in concert with a more general review of the four general education requirements and the processes used to assess them, with the aim of identifying a sustainable approach to implementing and assessing outcomes associated with the curricular requirements the university has determined to be important to the education of each of its students.

2. Offer continued opportunities for selected faculty to engage in aspects of this pilot process. Regardless of the decisions to be made in a full implementation of direct assessment measures for the competencies associated with the diversity curriculum, the positive outcomes of the pilot process suggest that similar activities for additional D1 and D2 faculty may constitute a valuable professional development experience and an opportunity to engage in close examination of the university's diversity curriculum. As such, we recommend exploration of an initiative to engage interested faculty in assessing student learning outcomes and their interaction with pedagogy and classroom environment, perhaps in connection with the new Diversity Faculty Fellows program. In particular, the norming activities and discussions that emerged in preparation for and during Rating Day could be replicated as a form of professional development that can support faculty members teaching D1 courses in communicating with one another and improving the design and implement their curriculum. D1 faculty involved in Rating Day described it as one of the most valuable meetings they ever had about their D1 courses, and the most honest in terms of how they perceived expectations, classroom climate, assignment design, and assessment of student learning outcomes. Our experiences suggest that these types of activities offer great benefit to faculty and warrant further consideration in conjunction with the university's ongoing commitment to the diversity curriculum and assessment activities associated with all areas within the General Education Requirement.

3. Review and modify institutional structures that support the D1 and D2 initiatives. The GEDAC is appreciative of the development of the General Education Requirement website that appears likely to enhance knowledge of the requirements and its associated processes, and recommends that the university continue to explore ways to enhance faculty understanding of the requirement, its associated competencies, and processes for designing and submitting courses for approval in the D1 and D2 categories. These institutional structures and processes may also serve to enhance systems level assessment of the diversity curriculum requirement. One major structure is the Course Leaf format, which some of the faculty described as an important tool that was not universally

understood by faculty submitting D1 and D2 materials. Additionally, if it were easier to view data collected over time through the Course Leaf tool, the GEDAC and DCRC would have more resources available to support and coordinate the work of the two committees, including important information regarding the degree to which all 14 competencies are being addressed across D1 and D2 courses.

4. Identify appropriate forums and committees in which to continue conversations about the ultimate expectations around the diversity requirement and the content and format of the competencies. The GEDAC and Rating Day participants engaged in multiple conversations that ended with questions about the degree to which there is university-wide consensus on the meaning of the diversity competencies and/or expectations for related student outcomes. The sheer number of competencies was noted as a barrier for promoting consistent faculty use as well as effective assessment of student learning outcomes. The GEDAC committee notes that discussions of the content and format of the Diversity Competencies have been raised in multiple forums, and we support additional review and potential revisions to the competencies. These points were also made in the AY2018 final report from the DCRC, in which the chair, Dr. Pablo Bose, noted several challenges in the review process of D1 and D2 courses. Most notably were the perceived “competing visions as to what the diversity requirement is expected to actually achieve at UVM” (p. 2). The GEDAC believes that the findings of our pilot study may be helpful to the DCRC as it moves to create clearer communication around the expectations for D1 and D2 courses, and refined review process. It could also create clearer guidelines for faculty who want to create new courses, as well as create important discussions about the challenges of identifying and assessing competencies selected for a given course.

5. Continue to explore opportunities to strengthen the diversity curriculum requirement as a part of the General Education Requirement, and to consider the future of assessment and oversight within the General Education Requirement. Although the GEDAC was not specifically charged to consider its place within the General Education Requirement, our experience across the year highlights the need to think about the future of the latter and its related assessment activities. We have acknowledged elsewhere that our work could not have proceeded as it did without the support of the Provost’s Office and the OIR. As we consider the future of the administration and assessment of each of the four components of the General Education Requirement, we note the resources that were available to us and hope that the university will engage in exploration of models that will best support and maintain the quality of this important work.