Association for the Assessment of Learning in Higher Education

AALHE

PROMOTING ASSESSMENT FOR LEARNING

INTERSECTION

A journal at the intersection of assessment and learning

Spring 2019

Contents

Note from The Editors-In-Chief
By Shannon Milligan and Catherine Wehlburg

Welcome to the 2019 Spring edition of AALHE's quarterly publication. The field of assessment has been in the news in many ways in the recent months – and not all of this was "good news." Certainly, the role of assessment in higher education is a crucial part of understanding student learning and of improving instruction. But what of the practices of assessment? How are assessment practitioners using data in ways to better inform decision making?

This issue of the *Intersection* provides a variety of ways that institutions are using data to better understand the process of education. Using a variety of assessment practices, the authors in this volume provide insight into the uses of assessment results to guide pedagogy and problem-solving. In some cases, data are already available and can be considered "found data" that might otherwise not have been used for improvement. These authors provide a peek into the work that they are doing and help to inspire us all to better consider our data and how to use what we have.

As has been our practice, the *Intersection* is foremost a publication about the practice of assessment, we continue to keep the voices of our authors as authentic as possible. We hope that you enjoy these articles and use them as an encouragement to meaningful, sustainable, valid, and reliable ways to better understand the assessment practices on your campus.

Letter from the AALHE President
By Jeremy Penn

Over the past few weeks here in Iowa the snow and ice has melted, the grass is beginning to turn green, and early-flowering trees are starting to produce small buds. I have begun contemplating plans for our small vegetable garden in the backyard of our home. Last year I had moderate success with cherry tomatoes – enough for some salads and salsa – but very limited success with the larger Early Girl tomatoes because they were devoured, prior to ripening, by a plump rabbit and several of his friends. The rabbits sat brazenly in the garden, evening after evening, munching away on a bounty of fresh vegetables, knowing that the holes they had chewed in the fence were large enough for them but too small for the dog. My efforts to prevent the rabbits from entering the garden with some fencing not only failed to keep the rabbits out, but actually provided the protection the rabbits needed to consume my garden at their leisure!

Yet, as I ponder which vegetables my rabbit friends will most appreciate this summer, I am struck with the recognition that my meager gardening efforts are only possible thanks to the labors of the many families who lived in this house before us. Prior owners hauled hundreds of large retaining wall blocks and hundreds of pounds of soil to form the raised garden. They had to carefully plan and manage the trees in the yard so sufficient sunlight would reach the garden. They had to feed the soil to make it suitable for vegetables by collecting food scraps and fertilizer and raking it into the soil. While I might like to think that my harvest of cherry tomatoes was due to my hard work and investment, the reality is that the crop I produced is the result of the combined efforts of many, over numerous years, who worked toward a common goal.

I reach a similar conclusion when I reflect on AALHE's status as an organization. AALHE is perhaps at its strongest point in its nearly 10-year history. We have a thriving conference, with 2019's conference in Saint Paul on-track to be the largest in AALHE's history. We have publications, such as this one, where members can share the latest in the work of assessment. We have the ASSESS Listserv, supported by the University of Kentucky, which continues to be a tremendously valuable resource for assessment professionals. We offer webinars, social media engagement, and a strong website for communicating our work to others. We have a full-time, paid Executive Director and are working on hiring some additional staff assistance. And, most importantly, we have volunteers, dozens and dozens of assessment experts and leaders who freely give their time to support AALHE. In just the last few months, volunteers provided hundreds of reviews for conference proposals, worked on a foundational statement describing what assessment is, participated in organizational planning meetings and served on committees, shared their expertise with others, and promoted the work of AALHE and of assessment in higher education through social media and multiple publications. I am thankful for these many contributions and hope you feel a sense of pride for all that you do and have done to build and support the work of AALHE.

Yet, we should not be too quick to pat ourselves on the back without acknowledging that our current success is largely due to the work of those who came before us. In its first decade, AALHE has been blessed with hard-working, thoughtful, caring volunteers who used their passion to build this organization. To them we owe a debt of gratitude.

At the same time, we should make use of this valuable opportunity to thoughtfully consider what we want AALHE to achieve in its second decade. AALHE is at an inflection point where our financial resources and organizational

strength enable us to turn our focus away from mere organizational survival toward the affect we want AALHE to have on the work of assessment, on our students, on our higher education institutions, and on our communities. How can AALHE best advocate for effective, meaningful assessment? What role does AALHE have in fighting for equitable, ethical assessment for all of our students? How do we prepare the next generation of assessment leaders? What partnerships should AALHE seek to support its mission? What research issues should be supported and pursued with energy and sustained effort? How do we make use of AALHE's resources to fulfill its mission? It will not be easy to answer these questions, and we may even disagree with each other, perhaps even strongly disagree with each other at times. However, it is only by seeking the answers to these questions together that we will recognize the potential that AALHE has to shape higher education in the years to come.

It is my hope, ten years from now, that we are celebrating ten more years of AALHE's contributions toward improving the work of assessment and improving student learning in higher education. I look forward to having a big party in 2029, where we come together to celebrate the many hours of hard work given by volunteers and the impact AALHE has had on the world. And if the rabbits are willing, I'll bring the salsa.

Association for the Assessment of Learning in Higher Education

AALHE

PROMOTING ASSESSMENT FOR LEARNING

Student Ratings of Instruction as "Found Data"
By Bruce Kelley

Abstract
Student ratings of instruction are used at almost every institution of higher education as a tool for faculty evaluation, and as a measure for annual review. The wealth of data that are generated by these surveys provides an assessment opportunity for a variety of units and departments. The University of South Dakota uses the IDEA student survey of instruction, and this survey has become "found data" for our Center for Teaching and Learning (CTL). This article examines how the CTL has used these data to help them better understand the effectiveness of various faculty development programs.

Student ratings of instruction are used at almost every institution of higher education as a tool for faculty evaluation, and as a measure for annual review. The wealth of data that are generated by these surveys provides an assessment opportunity for a variety of units and departments. The University of South Dakota uses the IDEA student survey of instruction, and this survey has become "found data" for our Center for Teaching and Learning (CTL).

Since CTLs often find themselves in the role of change agents (Ouellett, 2010, p. 9), and are, therefore, expected to facilitate the transformation of faculty instruction (Kelley, Cruz & Fire 2017, p. 4), we consider the IDEA evaluation results an essential piece of found data that allows us to triangulate results from a national database with other localized assessment tools to measure the effectiveness of our faculty development programs and initiatives.

For example, we use IDEA data to help us determine the effectiveness of our most intense faculty training experience, the Course Design Fellowship (CDF). This Fellowship is a one-week, 20-hour program that requires faculty to design (or redesign) one of their courses. The program is based on Fink's (2013) *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. This model starts with overarching goals—asking questions such as "how will students who take your class differ from students who don't?"—and moves to detailed plans for organizing content and structuring day-to-day learning activities. One of the goals of the CDF is to create courses that are active and engage students. We wish to encourage those outcomes and see improved IDEA scores for course design fellows, as measured two semesters before (F-2) and two semesters after the fellowship (F+2). Specifically, we track the raw IDEA converted average for the four primary scores on the IDEA summary sheet: Progress on Objectives, Excellent Teacher, Excellent Course, and Overall Summary. The scores are an average of the percentiles scored by each cohort.

Figure 1: Impact of Course Design Fellowship as measured by raw IDEA converted averages from 2 semesters prior to 2 semesters after the program
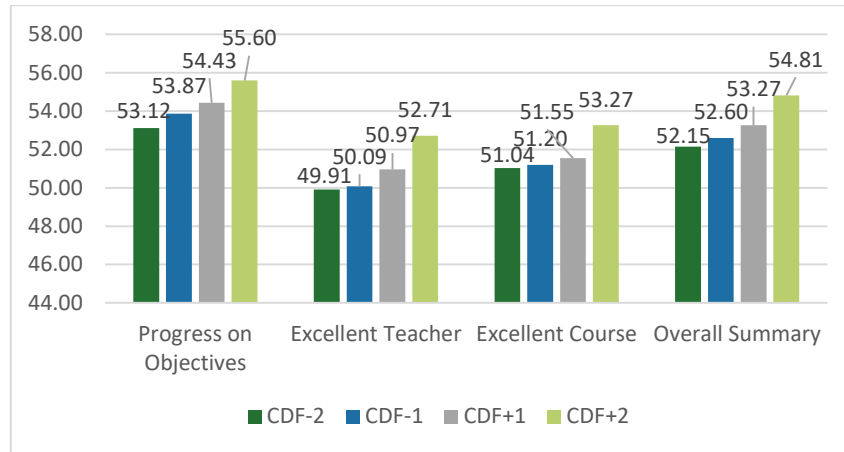
The program as a whole shows improvement in all scores between F-2 and F+2 as shown in Table 1.

Table 1: Average difference in raw IDEA converted averages between F-2 and F+2 for faculty who participated in CDF Fellowship AY11-AY16

	Value
Progress on Objectives	+2.47
Excellent Teacher	+2.80
Excellent Course	+2.23
Overall Summary	+2.66

We are able to triangulate these scores with our specific CDF assessments of students and faculty in the impacted courses. For example, the following percentage of students in the re/designed courses answered "often" or "very often" (on a 5-point Likert scale) on the following activities:

- Worked with other students on projects during class time: 92.2%
- The use of technology in this class enhanced my learning experience: 90.4%
- Engaged in course activities that helped you think in new ways: 81.2%

Students also indicated that these classes encouraged a supportive and helpful environment. 96.0% of students in re/designed courses indicated that they Strongly Agreed or Agreed with the statement that "Other students in this course are supportive," 96.1% indicated that they Strongly Agreed or Agreed with the statement that "Faculty members(s) teaching this class are available," and 94.3% indicated that "Faculty members(s) teaching this class are helpful."

Faculty are also surveyed, and the difference between their answers and the students is illuminating. For example, faculty indicated the following on their survey (each statement based on a 5-point Likert scale):

- Students worked "often" or "very often" with other students during class time: 56.3%
- Students used technology to meet the learning objectives for this course (agree/strongly agree): 93.8%
- The work in my course emphasized thinking critically and analytically (quite a bit/very much): 87.6%

Faculty often fear that implementing active learning will result in lower student ratings. Our data show that (on average) students appreciate the new kinds of learning that faculty require of them. In addition, the CDF assessment

provides qualitative evidence that students are engaged and participating in active learning, and that faculty efforts to enhance critical thinking are matched by students' experiences in the course. The IDEA data, combined with our survey of student learning activities and faculty observations, help us better understand the impact of our Couse Design Fellowship, both on faculty and on students.

Our CTL is also in charge of online faculty training and online course quality. Our goal is for online instructors who go through the quality assurance training process to have IDEA evaluations that match our face-to-face instruction. One way we measure this is by the percent of faculty who are at or above the national mean for the raw IDEA converted average for "Excellent Instructor" and "Excellent Course." Figures 2 and 3 show some of our more recent data:

Figure 2: Percent of faculty who are at or above the IDEA mean national average for "Excellent Instructor"
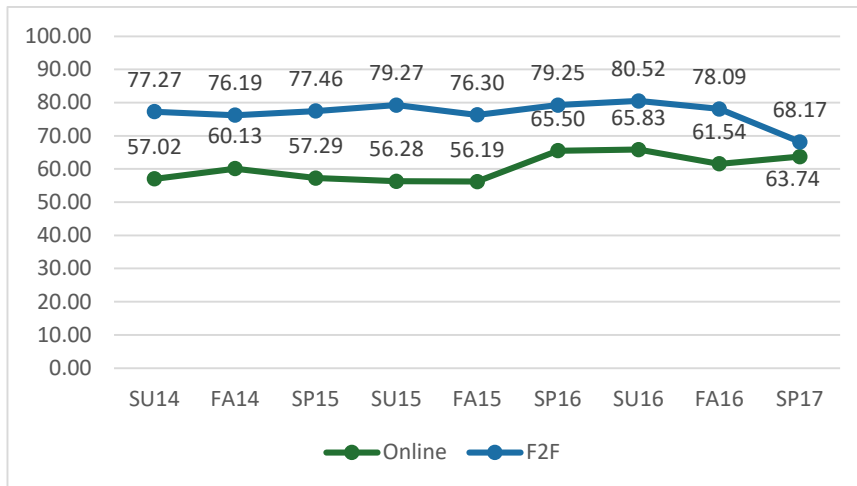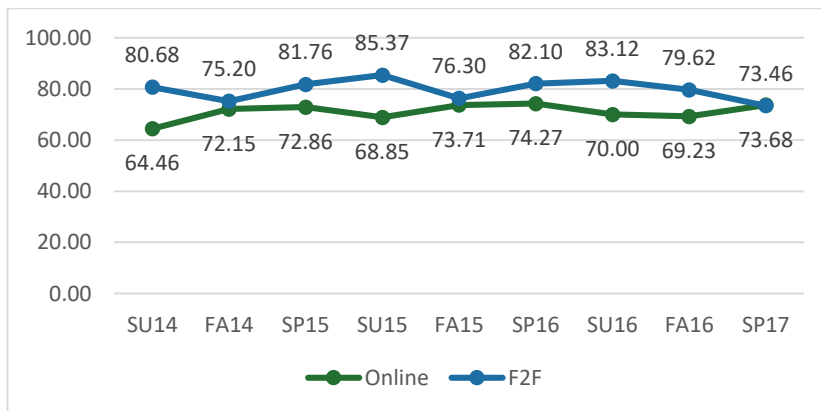

Figure 3: Percent of faculty who are at or above the IDEA mean national average for "Excellent Course"

It is a point of pride for the university that so many of our courses and instructors are highly rated. IDEA states that when these scores exceed 60%, the inference is that instructional effectiveness is "unusually high." Both online and face-to-face instruction meet this distinction. In addition, the gap between online and face-to-face courses has narrowed significantly for Excellent Instructor and has disappeared for Excellent Course. That is due in part to continued improvement in our online training process. However, there was also an unexpected dip in the face-to-face scores. This corresponds to a campus change to the online version of IDEA (Campus Labs) during the spring semester of 2017. This changed the data set, the way students submit evaluations, and the response rates. This highlights one of the challenges of using "found data"—sometimes decisions beyond your control impact the source or the nature of the data.

We triangulate this IDEA data with data from our Online Faculty Survey (OFS) to develop a more complete understanding of our training effectiveness. This survey is submitted to all online faculty during the Spring semester. Figures 4 and 5 show data from two of the questions on the OFS:

Figure 4: Percentage of total respondents who rated the CTL as helpful or very helpful on the OFS:
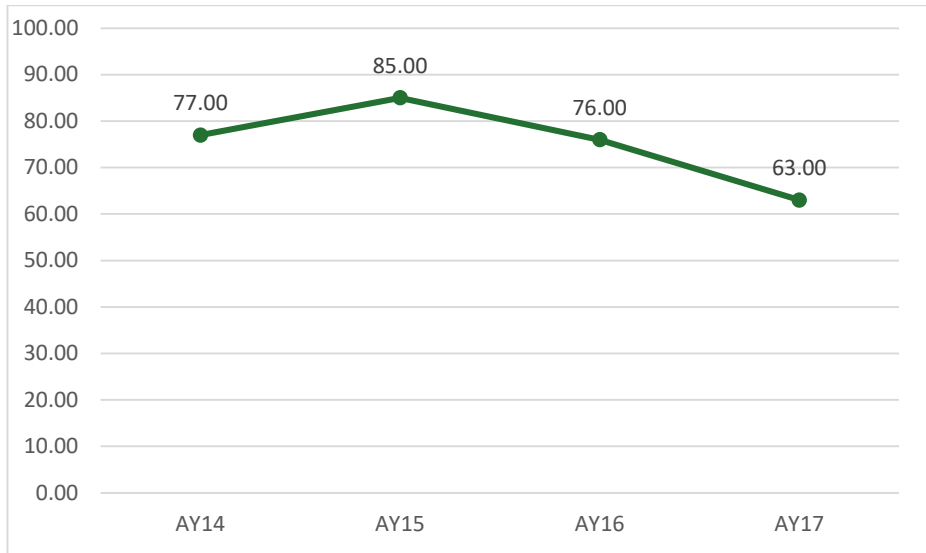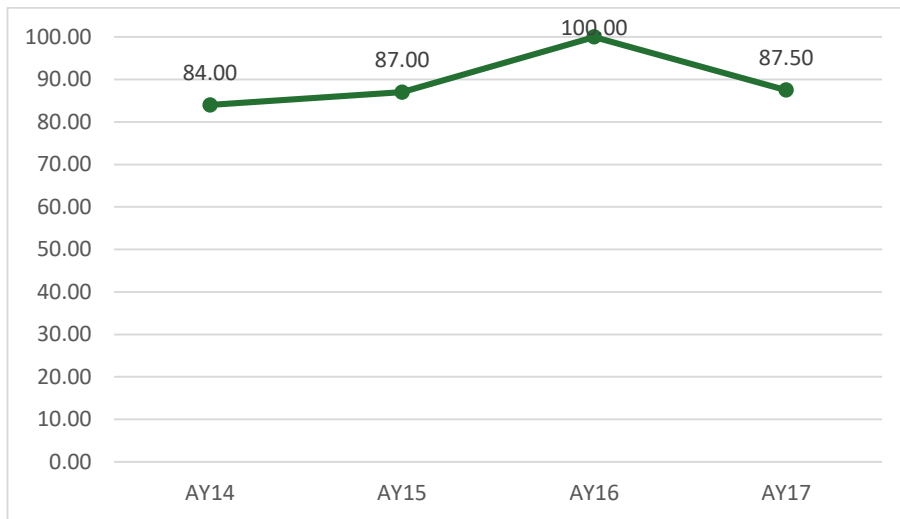

Figure 5: Percentage of new online instructors who rated the CTL as very helpful or helpful on the OFS:


We are obviously more pleased with the results in Figure 5, but effective assessment should identify areas where improvement is needed. Using the data from Figures 2-5, we can build the case that our training for online instructors is effective. Data from Figure 4 (and other related data sources) however, have also convinced me of the need to focus on customer service training both during our center's annual strategic retreat and during our staff development sessions this fall.

One of our greatest challenges each year is to effectively train last-second additions to our online faculty. While some late additions are expected, we also know that some of these late adds are due to simple procrastination by administrators in naming or hiring instructors for the sections, or procrastination by instructors in signing up for training. We used the IDEA data to help chairs and faculty understand the cost in student satisfaction that late training creates. We separated our online faculty into three groups—those faculty who were trained three weeks or

more prior to the start of classes, faculty who were trained 1-2 weeks prior to the start of classes, and faculty who were trained less than 1 week before classes start (Figure 6).

Figure 6: Percentage of courses that are at or above 3.5 on the IDEA ratings for "Excellent Instructor," differentiated by days between training and the first day of course.
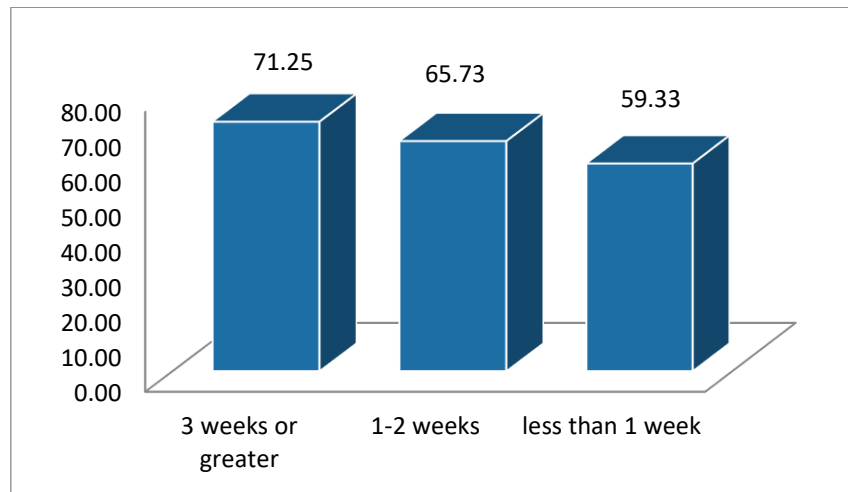
We found that there was a distinct impact on IDEA ratings based on when a faculty member was trained. What was perhaps most sobering for us is that these are career IDEA averages, and not just the semester of first teaching. In other words, faculty who are trained less than a week before classes begin perform relatively poorly throughout their online teaching career. While we hypothesize a number of reasons for the decline, IDEA data clearly support the importance of early training for online faculty. This year we plan to triangulate this data with DFW rates in these classes, perhaps linking student satisfaction directly with enrollment (or unenrollment) and student success.

Summary

Assessment has become a vital activity for all areas within higher education. Given limited resources (both money and time), we must use existing data in new and innovative ways. Institutions may have a treasure trove of data at their disposal by simply mining student surveys of instruction. These examples show how the CTL at the University of South Dakota has used these evaluations to assess professional development programs, improve training for online faculty, and advocate for early online training. Even if one doesn't use a commercial course evaluation system such as IDEA, it is quite likely that the data generated by your course evaluations can be used to inform teaching and learning activities across your campus. As such, "found data" can be an essential part of the campus assessment process.

References

Fink, L. D. 2013. *Creating Significant Learning Experiences*. San Francisco, CA: Jossey-Bass.

Kelley, B., Cruz, L. & Fire, N. (2017). "Moving Toward the Center: The Integration of Educational Development in an Era of Historic Change in Higher Education." *To Improve the Academy, 36*(1).

Ouellett, M. L. (2010). Overview of Faculty Development: History and Choices. In Gillespie, K. J. & Robertson, D. L. (eds.) *A Guide to Faculty Development*, 2nd edition. San Francisco, CA: Jossey-Bass.

Bruce Kelley, Ph.D., is the Assistant Provost at the University of South Dakota and can be reached at Bruce.Kelley@usd.edu .

Summary Circles: A Collective Auto-Ethnographic Approach to Assessing Learning Outcomes

By Dannielle Joy Davis, Farah Habli, Jamie Dehan, James A. Barker, Maureen M. Bell-Werner, Deborah R. Bush-Munson, Tyler A. De Shon, Alex Fajardo, Christina Luebbert, and Jeremy W. Main.

Abstract
Summary Circles refer to the pedagogical exercise of allowing students to divide course material and informally present information to the larger group for shared growth in knowledge. Utilized as an effort to promote engagement with material, expand knowledge, and encourage development of synthesis and presentation skills needed within higher education, the practice of Summary Circles may serve as a key in a faculty member's toolbox of educational teaching and learning strategies. The purpose of this work centers upon exploring Summary Circles as a tool in facilitating learning and assessment.

Introduction

Student involvement in teaching and learning practices set the educational platform to create and develop new instructional strategies. These strategies or approaches actively involve students in using learning materials to acquire deeper understanding. The Summary Circle is one of these instructional practices which bring out the best in students. This featured work illustrates implementation of Summary Circles in the graduate classroom, subsequent learning outcomes, and faculty course assessment. The piece takes a collective auto-ethnographic approach to understanding student and faculty perceptions of the learning strategy in cultivating skills needed for work in postsecondary settings, specifically administrative and faculty roles. Through collective auto-ethnography, this work offered students the opportunity to reflect upon course content as it related to their experiences (Davis et al, 2014). This research features the reflective writing of a faculty member and graduate students enrolled in the doctoral level course "The History of American Higher Education." Learning objectives for this course include:
1. To more fully understand and navigate the history of higher education.
2. To utilize and apply tools in conducting historical inquiry as it relates to current day policy and practice.
3. To develop transferable skills that may be applied as future administrators and faculty in higher education.

The professor and graduate students used journaling to reflect upon a core teaching method employed within the course, what the faculty member refers to as "Summary Circles." Exploring the utility of Summary Circles for training students as future faculty and staff in higher education serves as the primary intent of this work. The following reviews relevant literature related to instructional and learning strategies in higher education settings.

Review of Literature
In exploring the landscape of relevant instructional strategies that increase student learning, Eysink et al (2009) analyzed instructional approaches and compared their learning effectiveness, efficiency, and whether they resulted in different types of knowledge. The four instructional approaches included: (a) hypermedia learning (b) observational learning, (c) self-explanation based learning, and (d) inquiry learning (2009). Findings suggest that approaches prompting students towards self-explanation based learning and conducting experiments proved more effective and efficient than approaches that did not feature these learning methods (2009). The study further found participants acquired high-quality conceptual, procedural, and situational knowledge via such learning (2009).

Rather than reviewing varied approaches, Bjerkaker (2014) focused upon study circles which "at their best …offer learning without humiliation, learning without guilt and bad consciousness, non-violent learning, a humble way of learning, and learning for self-confidence and self- esteem" (p. 261, 2014). According to Bjerkaker (2014), the study circles learning method may best be described as: "learning by sharing," relying on each member's experience" (p.

261, 2014). This learning strategy offered experiences and opportunities for all to democratically participate using their prior knowledge and experiences (2014).

Overview of Summary Circles
Summary Circles comprise the active learning strategy of having each student select upon a variety of course related articles and book chapters. Students then are given one week to summarize, synthesize, and present material to classmates, thereby prompting them to become producers of knowledge for the course topic. They are given the option to present either formally (using Power Point or another presentation medium) or informally (sitting at a desk and verbally sharing material). Students then reflected upon their experiences with the Summary Circle exercise in individual course journals.

Significance of the Study
The results of this work point to the effectiveness of Summary Circles in the graduate classroom. Identifying an additional successful tool to be used in graduate level learning broadens available teaching strategies faculty may use in classrooms and aids in the ability to reach students with various learning styles. Development of additional presentation and synthesis skills through the use of Summary Circles aids in the professional cultivation of individual students, allowing for greater success in future endeavors as faculty and staff in higher education.

Methods
This work takes a collective auto-ethnographic approach (Lapadat, 2009) to the scholarship of teaching and learning. Researchers reflected on their lived learning and teaching experiences via writing (Delmont, 2009; Ellis & Bochner, 2000). According to Schwandt (2000) auto-ethnography "commonly refers to a particular form of writing that seeks to unite ethnographic (looking outward at a world beyond one's own) and autobiographical (gazing inward for a story of one's self) intentions" (p.13). The auto-ethnographic nature of this work requires collaboration between myself as a teacher-scholar and participating students. Hence, students who participated serve as the work's second authors. Using this auto-ethnographic approach, students employed journaling to share their experiences in the classroom with Summary Circles. The group collectively worked to decide upon journaling prompts, or guided questions, to use for their focused journal entries following several weeks of Summary Circle engagement.

Major Question
This work focuses upon the following major research question: How does the use of Summary Circles influence learning outcomes within the graduate classroom?

Data Analysis
Names were removed from journal entries and the constant comparative method was used to analyze data. With roots in grounded theory, the constant comparative method of data analysis is commonly used in varied forms of qualitative research and prompts the researcher to identify reoccurring patterns in the data (Merriam, 2009). Multiple researchers were used during the data analysis process via category and theme development to facilitate triangulation.

Participants
Participants consisted of a purposeful sample of doctoral students enrolled in the graduate course, The History of American Higher Education. In terms of gender, 8 women and 4 men took part in the work. Students self-identified in terms of ethnicity in the following ways: 2 of African descent, 1 of Middle Eastern decent, 7 as White, and 1 as Asian. Nine student participants served as administrators in higher education, with one serving in a faculty capacity. All participants were then current doctoral students.

Results
The following themes emerged as a result of the constant comparative method. These overarching themes are followed by student perspectives in light of the faculty member's pedagogical intent of training higher education faculty and administrators.

Professional Skill Development

Skills students believed were honed as a result of Summary Circles in the classroom included collaboration, synthesizing material, and public speaking. For instance one student noted:

> With Summary Circles, I am not only responsible to myself, but also to others…and thus I feel a greater sense of responsibility to grasp the main points of the text so I can best relay that to my peers. (Participant Reflection)

Other students shared how Summary Circles cultivated their critical thinking skills:

> I'm always intrigued and thankful when they [fellow students] summarize or emphasize something differently than I did, as it oftentimes exposes some of my own assumptions or questions that I brought into the readings. It's good to 'see' the readings through someone else and allow that to challenge some of my own assumptions and conclusions. (Participant Reflection)
>
> Summary Circles pushed my critical thinking skills as I consider information to pass on to my classmates. I can use this skill as a future teacher…(It also) helped me learn the material without getting bogged down. (Participant Reflection)

Students journaling about exercising critical thinking skills reflects the pedagogical intent of the instructor, while being student friendly via a decrease in reading load. Summary Circles further offer students the reoccurring opportunity of honing skills in synthesizing material and presenting it professionally to others: Common components of work for both faculty and administrators in higher education.

The Utility of Summary Circles

Students suggested that the utility of Summary Circles could be increased via student selection of course material. More time for discussion of Summary Circle articles/readings was also recommended. While the faculty member sees the benefit of expanding time for discussion, the need to present students with specific scholarly works as they relate to course content limits the degree to which students might select readings. One option providing a middle ground includes providing students specific guidelines in terms of topics to be selected for future Summary Circle sessions. These guidelines may be as broad as a list of topics to guide student selection of peer reviewed readings, to offering students opportunities to select from a set of specific articles and book chapters preselected by the professor.

Student Perspectives in Light of Faculty Pedagogical Intent

The faculty member's overall intent for the Summary Circles centers upon cultivating skills needed within the field of higher education both in terms of administrative and faculty work. These include summarizing and synthesizing content, presentation skills, and the ability to facilitate discussion. One participant noted:

> Summary Circles help me connect the past, present, and future. When I read, I look for similar circumstances that I encounter these days and how we can either prevent or sustain them in the future. As the saying goes, "Those who fail to learn history are doomed to repeat it." (Participant Reflection)

Perspectives that emerged from participants were sometimes challenged by the Summary Circle process. This meets the faculty member's objective of training faculty and staff in higher education to become critical thinkers and ethical decision makers. While some students experienced anxiety while preparing for or presenting during Summary Circles, the professor viewed this as a normal part of the learning process.

Faculty Assessment

Summary Circles center upon summarizing content, cultivating presentation skills, and honing the ability to facilitate discussion. These skills are needed within the field of higher education both in terms of administrative and faculty work. Based upon the faculty member's written and observed assessments, weekly Summary Circle activity resulted in all students demonstrating improvements in synthesizing course literature and oral presentation skills. Summary Circles further assisted in meeting the course objective of understanding and navigating the history of higher education, the topic of the featured course, through the skills of summary, synthesis, presentation, and discussion leadership. Exercising these skills meet the secondary course objective of exercising transferable skills for future

administration and faculty in higher education. This mirrors the work of Eysink et al (2009), as Summary Circles also prompted students to acquire high-quality conceptual and situational knowledge.

While on the surface Bjerkaker's description of study circles (2014) and this work's featured strategy of Summary Circles may appear similar, defining differences of the latter include its focus upon training graduate students to synthesize information and lead discussions based upon Summary Circle content. The additional aspect of journaling in this research offered a form of assessing student learning and revealed ways learning outcomes might be improved. In addition, student reflective journaling "point(ed) to evidence…that demonstrate(d) something positive resulted" due to faculty intervention (Welsh, 2018, pg. 9).

Conclusion

Future research in this area might further explore the demographics of participants and its possible connection to differing perspectives on learning strategies. Doing so may reveal generational, gender-based, or racial differences in respondents' views. Future work might also compare the experiences of classes using the Summary Circle method and the traditional lecture format. Recommendations for future teaching practice include continued use of the Summary Circle method as a form of active learning in both undergraduate and graduate education settings and the use of student journaling as a form of course assessment.

References

Bjerkaker, S. (2014). Changing communities. The study circle – for learning and democracy. *Procedia-Social and Behavioral Sciences, 142*, 260-267.

Davis, D.J., Coffee, T., Murphy, J., & Woods, J. (2014). Journaling our experiences: An auto-ethnographic approach to understanding graduate learning. *Reflective Practice*. DOI:10.1080/14623943.2014.944124.

Ellis, C. S., & Bochner, A. (2000). Auto-ethnography, personal narrative, reflexivity: Researcher as ubject. In N. Denzin, & Y. Lincoln (Eds.), *The Handbook of Qualitative Research* (pp. 733–768). Thousand Oaks, CA: Sage Publications.

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46,* 1107–1149. doi:10.3102/ 0002831209340235

Johnson, B. & Welsh, J. (2018). Interview with Dr. Barbara Johnson, Vice President at the Higher Learning Commission. *Intersection*.

Lapadat, J.C. (2009). Writing our way into shared understanding: Collaborative autobiographical writing in the qualitative methods class. *Qualitative Inquiry, 15,* 955-979.

Merriam, S.B. (2009). Qualitative research: A guide to design and implementation. San Francisco, CA. Jossey-Bass.

Schwandt, T. A. (2000). Dictionary of qualitative inquiry. Thousand Oaks, CA: Sage Publications.

Dannielle Joy Davis is a tenured Associate Professor of Higher Education at Saint Louis University and can be reached at djdavis@slu.edu.

Farah Habli is a doctoral candidate in Higher Education at Saint Louis University's School of Education and can be reached at farah.habli@slu.edu.

Jamie Dehan is a tenured Associate Professor and Program Director for the Stefani Doctor of Physical Therapy Program at the University of Saint Mary in Leavenworth, Kansas. She can be reached at jamie.dehan@slu.edu.

James Barker is a doctoral candidate at Saint Louis University in Higher Education and a Spanish teacher at Moberly High School. He can be reached at james.barker@slu.edu.

Maureen Bell-Werner is a Postdoctoral Fellow with Saint Louis University's School of Education and can be reached at bellwernermm@slu.edu.

Deborah Bush-Munson is a doctoral candidate at Saint Louis University and Program Director of the St. Louis Community College's Dental Assisting Program. She can be reached at deborah.bushmunson@slu.edu.

Tyler DeShon is an alumni of Saint Louis University's School of Education and can be reached at tyler.deshon@slu.edu.

Alex Fajardo is a doctoral candidate with Saint Louis University's Higher Education program and Adjunct Professor of Mathematics and Statistics. He can be reached at alex.fajardo@slu.edu.

Christina Tisher is a Mental Health Counselor at Saint Louis University and can be reached at christina.tisher@health.slu.edu.

Christine Luebbert is Assistant Director of the Ability Exhibit at Saint Louis University and can be reached at christine.luebbert@slu.edu.

Jeremy Main is Director of Field Education at Covenant Theological Seminary and can be reached at jeremy.main@covenantseminary.edu.

Using NSSE for Student Learning Outcomes Assessment: Validation through Expert-Panel Judgment

By Yao Zhang Hill

Abstract

This paper describes the process of justifying, acquiring, and validating the use of National Survey of Student Engagement (NSSE) items and data as indirect evidence of learning at a large public university. Specifically, it describes the use of expert-panel judgment procedure in identifying relevant NSSE items appropriate for assessing learning outcomes. The procedure enhanced content validity of this data source and increased faculty trust in the data, which further facilitated faculty's discussion and use of the assessment results.
.

What is NSSE?

National Survey of Student Engagement (NSSE) is a national survey commonly used by four-year colleges and universities to gauge undergraduate students' engagement. University of Hawaiʻi at Mānoa (UH Mānoa) administered NSSE in 2015 and received a record-high response rate (32%). Over 1300 undergraduate seniors responded to the survey and the sample represented the undergraduate senior population in general. Because of this, UH Mānoa Assessment Office considered the 2015 NSSE data as a valuable source of indirect evidence of student learning.

Why is NSSE a good choice for UH Mānoa to use as assessment data?

1. **Student self-reported efforts and activities on NSSE has the potential to serve as indirect evidence of learning.** To assess learning, evidence can be collected using direct and indirect methods. Direct evidence of learning is often actual student work (e.g., written reports) and indirect evidence "capture students' perception of their learning" (Maki, 2004, p. 88).

   NSSE in general consists of two types of questions: (1) student efforts and activities, and (2) institutional resources and learning opportunities. Items that capture students' self-report on academic efforts, activities and behaviors are valuable as indirect evidence of learning because those reflect students' knowledge, skills, and dispositions. For example, an NSSE item like "2.b *How often have you prepared two or more drafts of a paper or assignments before turning it in?*" refers to students' writing habits which indirectly reflect their revision skills in writing. On the other hand, institutional resource questions such as "*How much does your institution emphasize the following?* (NSSE 2015 Item 14)" speak to institution's intention and effort, not students' value and actions, thus are not appropriate to serve as evidence of learning.

2. **Using NSSE as indirect evidence of learning is supported by psychometric analysis and used in practice**. Zilvinskis, Masseria, and Pike (2017) used canonical correlation analysis and found that NSSE scalelets (sets of questions assessing the same concept) were related to students' self-reported learning outcomes. For example, *higher order and reflection learning* scalelet was positively related to self-reported *academic* and *interpersonal learning* outcomes. This study provided some concurrent validity evidence for using NSSE, or at least some NSSE items, as indirect evidence of learning.

   In practice, some colleges and universities have used relevant NSSE items to report the results on their institutional learning outcomes. Eastern Illinois University looked at the questions that speak to undergraduate learning goals, summarized the results of the items under each learning objective, and shared the results with the assessment committee (personal communication with Karla Sanders). At Miami Dade College, they mapped particular items to their ten college-wide learning outcomes and produced reports across the institution, as well as for each campus (personal communication with Marisa Yates).

3. **NSSE 2015 data has good quality and fills the gaps in institutional assessment results**. Many psychometric studies were done to validate NSSE's psychometric constructs using national datasets (e.g., Carle, et al., 2009; Zilvinskies et al., 2017). To be confident in our own data interpretation, we collaborated with Jessica Miranda, a researcher in the College of Education, and investigated the reliability and psychometric quality of the survey items with UH Mānoa datasets using the item response theory. We only found two unreliable items and were satisfied with the quality of the rest of the items for assessment purposes.

    UH Mānoa has ten undergraduate Institutional Learning Objectives (ILOs) (http://go.hawaii.edu/KfG). We are also required to assess the five core competences specified by Western Association of Schools and Colleges Senior College and University Commission (WASC):
    - (1) critical thinking
    - (2) information literacy
    - (3) oral communication
    - (4) quantitative reasoning
    - (5) written communication

    As of the end of 2017, we have results on four out of five WASC Core Competencies, all from direct evidence, and none on disposition-based ILOs (e.g., respect for people & culture,). NSSE data have the potential to fill the gap in our institutional learning assessment results and can provide insightful information on student learning experience.

Aligning NSSE Items with ILOs: Content Validation through Expert Judgment

After we made the decision that NSSE data has sufficient quality and is an appropriate source of assessment evidence, the next step was to align relevant NSSE items with institutional-level learning outcomes. Often, this kind of item-outcome alignment work is done by assessment professionals, analysts, or institutional researchers because they are familiar with both the survey and the outcomes. The drawbacks are:
- (1) It is possible for other campus stakeholders to interpret the survey items differently because of their content expertise.
- (2) If faculty and staff are not educated on the survey and the outcomes, they would have a harder time interpreting and using the results.

UH Mānoa Assessment Office used expert-panel judgement to select the outcome-aligned NSSE items. We designed this process as a content validation procedure, also as a professional development opportunity for campus stakeholders. The item-outcome alignment using expert-panel judgement takes three steps:
- (1) identify key expert groups to serve as the panel
- (2) consolidate group responses
- (3) discuss and finalize alignment

Step 1: identify key expert groups. The groups need to consist faculty, staff, and administrators who have fundamental understanding of assessment and institutional-level learning outcomes, as well as content area expertise (e.g., civic engagement, quantitative reasoning). It would be desirable for the members to be in a position to influence change. The first group we identified was Institutional Learning Objectives Implementation Committee (ILOIC), consisting of about 10 active members, who are faculty leaders on academic initiatives, an assessment specialist, directors of curricular (e.g., General Education) and co-curricular (e.g., service-learning) programs, as well as the Associate Vice Chancellor for Academic Affairs and the accreditation liaison. We also identified campus Assessment Committee (later became Educational Effectiveness Committee) and General Education Committee as potential groups.

At the end of January 2017, the ILOIC group received a description of the project and each individual members made judgement on how each NSSE item was aligned or not aligned with an institutional-level outcome. Prior to their individual work, we provided a short training on identifying survey items appropriate for assessing learning. Ten members submitted their work.

In **Step 2: consolidate group responses**, Assessment Office staff recorded each person's responses and tallied the results. The tally shows how many times a survey item was marked as related to a particular outcome. Figure 1. Illustrates part of the tally. For example, for Item 1.a (*How often have you asked questions or contributed to classroom discussions?*), no one considered it aligned with the *Written Communication* outcome, but four people considered it aligned with *Oral Communication*, and one considered it aligned with *Critical Thinking*.

Item	Written Communication	Oral Communication	Information Literacy	Critical Thinking	Quantitative Reasoning	…
1.a	0	4	0	1	0	…
1.b	7	0	0	0	1	…
1.c	0	0	0	0	0	…
1.d	0	0	1	0	0	…
…	…	…	…	…	…	…

Figure 1. Tally of NSSE Item Alignment with the Learning Outcomes

Item-outcome alignments identified by the majority (6 or more members) were endorsed as the item being appropriate for assessing that aligned outcome (e.g., Item 1.b for assessing *Written Communication*)

In **Step 3: discussion and finalization**, we brought the item-outcome alignment results back to the ILOIC, led group discussion on controversial items, and helped the group reach consensus. Aligning NSSE items through expert-panel judgment allowed us obtain assessment results on all five WASC Core Competencies and six out of ten ILOs, namely:

a) Thinking critically and creatively
b) Conduct research
c) Communication and report
d) Continuous learning and personal growth
e) Respect for people and culture
f) Civic participation

After faculty made the alignment decisions, they discussed and selected survey responses (e.g., *never*, *sometimes*, *often*, *very often*) that indicated the acceptable level of learning. For most items, the top two options (e.g, *often* and *very often*, *quite a bit* and *very much*) represented frequent participation or substantial learning deemed as acceptable level of learning for our students. We made individual judgements on questions such as number of papers written (3 or more were considered satisfactory as evidence of writing ability) and number of hours doing community service or volunteer work (1 hour or above were deemed acceptable as evidence of civic participation).

Analyzing and Using NSSE Data for Learning Assessment

After we finalized the item-outcome alignment and decided on acceptable responses as evidence of learning, we calculated the percentage of the students who demonstrated frequent participation or substantial learning on each relevant item. This becomes the item percentage score. The outcome score was calculated as the average percentage score of the items aligned with that outcome. We used the data collected from approximately 1300 college seniors to represent student learning at the exit level.

In addition, we indicated how UH Mānoa NSSE item scores compared to the peer institution averages, using data reports provided by NSSE. The peer institution comparison was useful especially when NSSE results contradicted our direct assessment results. For example, even though only 55% of the sampled seniors demonstrated substantial

learning on NSSE in written communication, most of the item scores were significantly higher than peer institution averages. This is congruent with our direct assessment finding that written communication has been the highest achieved WASC Core Competencies at UH Mānoa.

We summarized and presented the results in different formats to different campus groups. The general outcome summary results were used for General Education program review, WASC accreditation, and to general campus groups (e.g., ILOIC, General Education Committee, Assessment Committee). When presenting to groups in specific curricular areas (e.g., Librarians, Oral/Written Communication General Education Committee, General Education Foundation Committee), we provide more detailed results, including diagnostic information on our students' strengths and weaknesses.

In all data reports, we included both NSSE results and results from direct assessment projects when available. To help faculty digest the results, we sent out the results several days prior to the meeting with narratives of key findings. When there was a chance to present the results in a face-to-face meeting, we strategically used most of the meeting time to lead discussions on ways to use results, with only brief presentation of the results.

For instance, when presenting the results to the ILOIC group, we encouraged the group to prioritize learning outcomes to act upon for improvement purposes. We provided guiding principles to prioritize outcomes for action:
> "*Consider acting upon:*
> *(1) The lowest achieving outcomes;*
> *(2) Where resources are available for effective actions—to make sure that we have the capacity to change; and/or*
> *(3) Where we have movers and shakers in the area to make the change –to leverage existing momentum for the change to happen.*"

We facilitated the use-of-results discussion and used the dot-voting technique to select three priorities for the committee to act upon assessment results: quantitative reasoning, information literacy, and civic engagement (see dot-voting results in Figure 2). In the subsequent meetings, the ILOIC committee detailed their action plan.
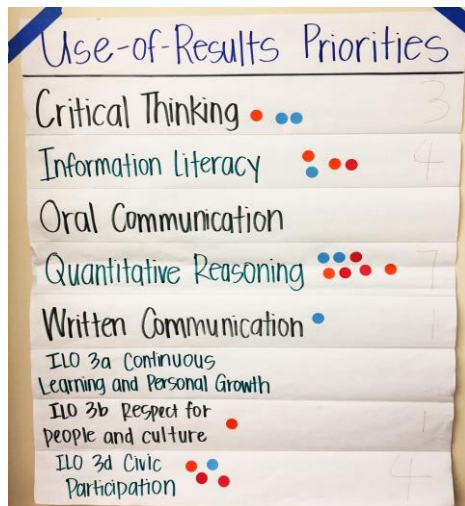

Figure 2. Dot-Voting Use-of-Results Priorities

NSSE results enriched the campus understanding of students learning and stimulated conversations and decisions to improve learning. For example, in the area of information literacy, we understood that students frequently used outside sources but many often did not engage in critical evaluation of sources, indicated by several NSSE items related to rejecting sources of questionable quality, identifying contribution of the source to the field, and using information to accomplish a purpose. This is an action-ready result. When we presented this result to the librarians and the Written Communication Committee, both groups came up with ways to act upon the results. Many librarians were motivated to seek out faculty collaboration in an effort to provide increased training to their students. The

Written Communication Committee began to discuss how to include information literacy as part of the requirement when faculty apply for the writing intensive course designation.

Conclusion

Evidently, using relevant NSSE items as indirect learning evidence immediately facilitated the assessment of institutional level learning outcomes and facilitated accreditation self-study. It provided results on six outcomes with no prior data sources. It complemented the existing findings related to the four outcomes previous assessed.

At a deeper level, our NSSE work promoted campus discussions and the use of assessment results. Employing an expert-panel in identifying outcome-aligned NSSE items is a process that honors campus stakeholders' input and expertise. The fact that we went through such a process increased campus groups' trust in the results because the relevance and appropriateness of the items have already been vested by their colleagues. This trust was enhanced by the fact that we verified the psychometric quality of our NSSE data and that NSSE results are congruent with our existing results from direct assessment. The trust in the relevance and quality of the NSSE results facilitated faculty's discussion and the use of the findings. NSSE items provided additional insights and diagnostic information into our institutional level learning outcomes.

To conclude, higher education institutions with NSSE data can leverage this data source as indirect evidence of learning. Using the expert-panel to make judgements on outcome-aligned NSSE items can greatly increase the relevance and utility of this data source.

References

Carle, A. C., Jaffee, D., Vaughan, N. W., & Eder, D. (2009). Psychometric Properties of Three New National Survey of Student Engagement Based Engagement Scales: An Item Response Theory Analysis. *Research in Higher Education*, *50*(8), 775–794. https://doi.org/10.1007/s11162-009-9141-z

Maki, P. (2004). *Assessing for learning: Building a sustainable commitment across the institution* (1st ed). Sterling, VA: Stylus.

Zilvinskis, J., Masseria, A. A., & Pike, G. R. (2017). Student Engagement and Student Learning: Examining the Convergent and Discriminant Validity of the Revised National Survey of Student Engagement. *Research in Higher Education*, *58*(8), 880–903. https://doi.org/10.1007/s11162-017-9450-6

Yao Zhang Hill is an Associate Specialist in the Assessment Office at University of Hawaiʻi at Mānoa and can be reached at yao.hill@hawaii.edu.

Supplemental Instruction: A Closer Look at the Data
By Regina Roof-Ray

Abstract

In order to evaluate the impact of voluntary test-specific supplemental instruction (SI) in Psychology 101, data were collected and compared between attenders and non-attenders. While those who attended SI sessions did have a better average test score, their first to second test score performance decreased more than those who did not attend. This, and other variables, illustrates several potential factors to consider in the use of voluntary test-specific supplemental instruction.

With the current focus on student retention in higher education, many articles have been written that promote student engagement as a means of reducing attrition (Breser, 2011; Kerby, 2007; Lynch, 2012). Supplemental instruction is one strategy that is suggested as a method of improving student learning (Blanc, DeBuhr, & Martin 2016; Moleko, Hlalele, & Mahlomaholo 2014). There have been many webinars and journal articles that present data demonstrating that SI is correlated with higher grades and student retention. For example, "The Effects of Supplemental Instruction on Timely Graduation" by Bowles, McCoy, and Bates and "Supplemental Instruction: Improving First-Year Student Success in High-risk Courses" by Stone and Jacobs, promotes supplemental instruction.   Indeed, it would appear to be hard to argue against the benefits of providing additional instruction to students to help them engage in a learning community, interact with material, and develop a support network. However, since every class may have its own unique culture of learning, it is important to assess the actual success of supplemental instruction with each specific class and evaluate what those results may actually mean.

The purpose of this article is the show how one community college assessed the effectiveness of SI in general psychology courses. This study looked at 94 students who attended SI within the context of a general psychology course and 71 students who did not use SI. All sections were taught by the same instructor. The results were gathered using both informal and formal data collection opportunities.

Eight semesters of data were collected in order to evaluate the impact of test-specific supplemental instruction in Psychology 101. A psychology tutor with a long, successful history of working with psychology faculty provided the supplemental instruction and all students were made aware of the free service. Two sessions were held before tests in order to accommodate more students. Attending students were awarded four extra credit points on their course cumulative point total (possible total course points was 650).

The collection of data began in Fall 2014. At this point, data collection was quite informal. Supplemental instruction was not offered before the first test during this initial semester. We found that the students who chose to attend SI before the second test had an average 18% increase in their scores from the first test to the second. At first glance, this would indicate success as all students took the same exam (the only difference was that the order of questions was modified). However, during the following semester, Spring 2015, the average increase from the first test to the second test was significantly less at only .7% even though they used SI. Interestingly, six of the fourteen students actually had a decrease and eight students had an increase after attending supplemental instruction. The increase from the first test to the third test, when attending SI, was an average of 5.4%. While, on the whole, this seems like SI had a positive result, there were several variables that could have impacted this data. First, only six students participated by taking advantage of SI in Fall 2014 and only fourteen in Spring 2015. Second, the data does

not indicate anything about the difficulty level of the first, second, or third test. Third, the data does not address the motivation of attenders for the SI.

To gain a better understanding of the effect of the extra instruction, in Fall 2015 a random sample of non-attenders was also tracked. This allowed for a comparison of attenders of SI to non-attenders. Again, at this point, the data collection was still informal, but the additional information helped address the potential variable of the difficulty of tests in order to provide a better picture of the success of SI in terms of impact on grades. On the first test, those who attended SI scored an average of six points higher than those who did not attend. It is possible, though, that this might indicate that those who attended SI were already doing better than those who did not attend SI. The second test was the highest scoring test for all students. This could indicate that the previous Fall's data, indicating better scores on the second test for those that had SI, may not have been a product of the SI. However, those who attended SI did score an average of ten points higher than those who did not attend on the second test and an average of nine points higher on the third test.

From Spring 2016 through Spring 2018 a similar approach to data collection was taken. All tests were worth 100 points and SI was available for all tests. In the Spring 2016 semester the average test score for those who attended SI was better than the average for those who did not (5 points on the first test, 7 points on the second, and 15 points on the third). Five students who did not attend SI for the first test but attended the second had an average increase of 1.1 points, while those who did not attend either session had an average decrease of 4 points. Four students who attended the SI for the second test but not the third test had an average decrease of 4.5 points while those who did not attend either session had an increase of 2 points. These results may address some of the concern as to whether those who were attending SI were more motivated students who would have been more successful even without SI.

Subsequent semesters showed similar data:

	Test 1 average with SI	Test 1 average without SI	Test 2 average with SI	Test 2 average without SI
Spring 2017	88	74	86	66
Fall 2017	86.9	79.4	82.2	68
Spring 2018	71.7	69.6	74	67.5

While the class averages varied from semester to semester, those who attended SI sessions consistently performed better than those who did not. In order to gain a better understanding, more formal data was collected in Fall 2018. Fifteen students who attended SI before two tests and fifteen students who did not were included in the data collection. Ultimately, 29.8% of eligible students attended SI for at least one of the tests. There were eight students who attended SI for the first but not the second test. Their data was analyzed separate from the two groups of fifteen.

Of those attending SI for both tests, the average first test score was 80.2% and the average second test score was 74.2%, while those who did not attend SI had an average first test score of 72.5% and average second test score of 69.9%. Average course grades for SI attendees was 3.03 (B) and for non-attendees was 2.65 (C). The students who attended SI for the first test but not the second had an average decrease of 4.9%; however, the average decrease for those who did attend for both tests was 6.0% and the average decrease for those who did not attend SI at all was 2.6%.

	Test One	Test Two	Difference	Average Course GPA
SI Attenders	80.2	74.2	6.0	3.03
Non-attenders	72.5	69.9	2.6	2.65
Attended for Test one but NOT for test two	77.3	72.4	4.9	2.75

2 sample T-test for Test One H_O $\mu_1 = \mu_2$ t = -2.06
H_A $\mu_1 < \mu_2$ p = 0.02
There was a significant difference at the α = .05 level

2 sample T-test for Test Two H_O $\mu_1 = \mu_2$ t = -1.17
H_A $\mu_1 = \mu_2$ p = 0.12

The test performance for those who attended SI was better on both tests for each semester. The average course GPA for those who attended SI was also better by an average of .38 points. Interestingly, there was less test decline for the non-attendees than there was for the attendees or those who attended for the first test but not for the second test. This would support the idea that those who were more likely to study in the first place were more likely to attend SI, but also that the SI may have provided a false sense of confidence, as demonstrated by first to second test decline in SI attendees. Additionally, the SI for the first test may have better prepared students for the structure of the test, while on the second test this would have been a known factor for all students.

Overall, there does seem to be some grade improvement benefit to those students who participated in SI. Consistent data collection and evaluation suggests, however, that it is not as significant as first believed in this population. While further research must be done, these findings suggest that the positive results we see related to test-specific SI in the general psychology classroom may not only be the result of the SI itself. It is possible that those who seek SI are more concerned about grades and understanding of material and would have put forth more effort to be successful with or without SI. It is also possible that the SI increased accountability through the group interaction of the attendees, as those who attended had the added qualitative benefit of engagement with other students and college resources. The self-fulfilling prophecy may also have played a part in these finding. On a less positive note, it is also possible that the SI provided a false sense of confidence, as demonstrated in the increased drop in grades from the first to second test of SI attendees. There are many variables that can impact student learning and SI can be an important addition to the teaching/learning process.

References

Blanc, R., DeBuhr, L., & Martin, D. (2016). Breaking the attrition cycle. *The Journal of Higher Education, 54*:1, 80-90. DOI: 10.1080/00221546.1983.11778153

Bowles, T., McCoy, A.B., & Bates, S. (2008). The effects of Supplemental Instruction on timely graduation. College Student Journal, 42(3), 853-859

Brewer, M. R. (2011). *Student success: A qualitative analysis of the engagement of the successful first-generation student* (Order No. 3450540). Available from Psychology Database. (864832977).

Dalton, C. (2011). *The effects of supplemental instruction on pass rates, academic performance, retention and persistence in community college developmental reading courses* (Order No. 3485052).

Kerby, M. B. (2007). *Measuring student engagement activity levels and persistence: A study of a four -year comprehensive university* (Order No. 3401265). Available from Education Database. (250910086).

Lynch Ervin, S. E. (2010). *Examining student engagement of African American community college students: Focusing on the CCSSE benchmarks* (Order No. 3419535). Available from Education Database. (753947095).

Moleko, M., Hlalele, D., & Mahlomaholo, M. (2014). Challenges experienced with the implementation of supplemental instruction at institutions of higher education. *Mediterranean Journal of Social Sciences*, 5(27).

Stone, M. E., & Jacobs, G. (Eds.). (2008). Supplemental instruction: Improving first-year student success in high-risk courses (Monograph No. 7, 3rd ed.). Columbia, SC: University of South Carolina, National Resource Center for The First-Year Experience and Students in Transition.

Regina Roof-Ray is an Assistant Professor of Psychology at Harford Community College and can be reached at rroofray@harford.edu.

CALL FOR PAPERS - deadline June 14, 2019

Summer 2019 Intersection Theme: *Equity in Assessment, Part 1*

Recently, conversations have arisen on a national level around the diversity of today's college students, and in particular how assessment of student learning needs to be reflective of and responsive to the needs of our diverse students. This **Summer 2019** edition of AALHE's *Intersection* will continue the dialogue on equity and diversity in assessment that will start at the June 2019 AALHE conference and lead up to the June 2020 conference. We are looking to present our readers with examples of successful strategies for increasing equity in assessment, including, but not limited to, implementation of culturally responsive assessment practices, means for increasing cultural competence in assessment, and strategies for discussing equity in assessment with campus assessment partners/stakeholders. Please consider sharing your successful assessment processes to assist your AALHE colleagues in this important endeavor.

Submissions should be sent by **JUNE 14th** to Editor-in-Chief Shannon Milligan, Ph.D. at shannon.milligan@gmail.com.

General Guidelines: Articles should be 700-2100 words and of general interest to assessment professionals. The emphasis is on readability over technical details, although the use of statistics is not discouraged. APA style should be used for citations. Submissions should not be used to market specific products.

★★★★★

Check out AALHE's New Podcast!!

Hosts Erin Milne and André Foisy explore current events, contemporary issues, helpful strategies for measuring student learning outcomes, and stories from the assessment community in this monthly podcast. We'll also feature interviews with assessment professionals from around the country.

Please send us your assessment success stories and challenges by e-mailing an audio file to podcast@aalhe.org or leaving a message at 1(302)8-ASSESS. We'd love to hear how assessment works (or doesn't) for you. Your audio may be used in an episode.

Join us for **AALHE 2019** at the InterContinental Hotel in Saint Paul, Minnesota

AALHE's Ninth Annual
Assessment Conference
Saint Paul
Minnesota
InterContinental Hotel
JUNE 10-13, 2019

Association for the Assessment of Learning in Higher Education

AALHE

PROMOTING ASSESSMENT FOR LEARNING

INTERSECTION