

Class observations from the University of Hawai‘i at Mānoa highlight the need for active learning strategies to support diverse students in large classes

Jennifer Engels¹, Barbara Bruno¹, Noelle Dasalla², Daniela Böttjer-Wilson³

¹Hawai‘i Institute of Geophysics and Planetology, University of Hawai‘i at Mānoa, United States of America, ²Division of Natural Sciences and Mathematics, Chaminade University, United States of America, ³Center for Teaching Excellence, University of Hawai‘i at Mānoa, United States of America.

Abstract

Compelling evidence indicates that “active learning” (learning by doing) is an effective pedagogy regardless of discipline or class size, and can be particularly effective with diverse students. This study investigated active learning practices in 64 classes at the University of Hawai‘i at Mānoa, a US university with a highly diverse student body, using a “Passivity Indicator” (PI: ratio of class time spent in passive activities to total class time). For all classes, the mean PI was 43%. Statistical analysis reveals no significant differences in the PI of classes taught in STEM vs. non-STEM disciplines, or between upper vs. lower division courses. However, the PI in larger classes was found to be significantly greater than in small classes (64% vs. 39%, respectively; $p=0.02$). Moreover, classroom activities aligned with an active learning standard in Language and Literacy Development (e.g., students answering questions) occurred twice as often in small (24%) vs. large classes (12%, with $p=0.02$). Altogether, these findings indicate an opportunity for more active learning in large classes. We present a range of research-based pedagogical strategies that can be readily implemented in large classrooms, and encourage instructors to use their implementation as research opportunities to gather data on student success.

Keywords: Classroom Observations, Active Learning; Diversity; Hawai‘i.

1. Introduction

Active learning (*i.e.*, “learning by doing” or “student-centered learning”) has been shown to improve subject retention, increase student engagement, and reduce failure rates (*e.g.*, Springer *et al.*, 1999; Ruiz-Primo *et al.*, 2011; Freeman *et al.*, 2014), particularly for women (*e.g.*, Lorenzo *et al.*, 2006) and minorities (*e.g.*, Haak *et al.*, 2011). The Classroom Observation Protocol for Undergraduate STEM (COPUS) quantifies the degree to which a course uses active learning practices. For each two-minute interval of class time, a trained observer classifies what students and instructors are doing using 25 activity codes (Smith *et al.*, 2013). A recent nationwide (USA) COPUS analysis of 2000+ University classes found that the majority (55%) were taught using a “didactic” style ($\geq 80\%$ lecturing); only 18% used a “student-centered style” ($\leq 50\%$ lecturing). Large classes were especially likely to be didactic (Stains *et al.*, 2018). In light of the overwhelming benefits of active learning, these results are troubling, particularly for minority-serving institutions like the University of Hawai‘i at Mānoa (UHM), one of the most diverse universities in the US (US News and World Report, 2019).

The Center for Research on Education, Diversity, and Excellence (CREDE) funded 31 research projects on teaching culturally and linguistically diverse students (Yamauchi *et al.*, 2016). The resulting CREDE standards are evidence-based best practices based on Vygotsky (1978) and 40+ years of research on interaction-rich dialogues to promote conceptual understanding (Tharp *et al.*, 2000). The US Department of Education ranked the CREDE standards the most effective for promoting reading achievement, and 2nd most effective for improving English language literacy among 73 studies of language development for English language learners (Yamauchi *et al.*, 2016).

This case study aims to: (A) assess the degree to which active learning is being used in UHM classrooms via COPUS; (B) identify any correlations between the level of active learning and certain class characteristics (class size, subject content, and academic level); and (C) evaluate the enactment of CREDE standards for teaching diverse students.

2. Methods

This study was conducted over six semesters (2015-18). After receiving UHM Institutional Research Board (IRB) approval to work with human subjects, COPUS observations were conducted by five observers, whose training included video and in-class practice. Inter-rater reliability exceeded 95% among all 5 observers using Jaccard (1901) similarity scores, indicating robustness. Each two-minute class interval was categorized as *Passive* (intervals during which students were only “listening” and/or “waiting”) or *Active* (intervals including at least one student activity other than listening or waiting) (Smith *et al.*, 2013). We define the “passivity indicator” (PI) as the ratio of passive intervals to total intervals.

To determine whether student activity levels correlate with certain class characteristics, each class was categorized in three ways: A) STEM (0) vs. non-STEM (1), B) Upper (0) vs. Lower (1) Division, and C) Large (0), *i.e.*, >50 students, or Small (1), *i.e.*, ≤50 students. (Binary numeric “dummy” variables (0 and 1) were assigned to enable multivariate analysis). For each category, the mean and standard deviation of PI were calculated. Using box plots and analysis of variance (ANOVA) type III analysis, PI were compared first within each class category, and then across class categories to check for interactions, using significance level $\alpha= 0.05$.

Lastly, COPUS codes were compared to three CREDE standards: Joint Productive Activity (JPA), Language & Literacy Development (LLD), and Instructional Conversation (IC). For each, COPUS codes were assigned corresponding to the class times during which the CREDE standards were or were not enacted (Table 1). Mean percentages of class times spent enacting each CREDE standard were then calculated for all classes.

Table 1. CREDE Standards Performance Continuum & corresponding COPUS Codes for evaluating the enactment of CREDE Standards in the classroom.

CREDE standard¹	Enacted (COPUS Codes) ^{2,3}	Not Enacted (COPUS Codes) ^{2,3}
JPA	Students and instructor collaborating in joint product (student: OG; instructor: MG)	Student working on individual products (student: Ind; instructor: W)
LLD	Instruction in which academic language use by student predominates (student: AnQ, SQ; instructor: AnQ, PQ)	Instruction dominated by instructor talk (student: L; instructor: Lec)
IC	Goal-directed, fully inclusive conversation between instructor and small group of students (student: - ; instructor: 1o1)	Informal, non-academic discourse in whole-class settings (student: WC; instructor: -)

¹JPA (Joint Productive Activity); LLD (Language & Literacy Development); IC (Instructional Conversation) (CREDE, 2019); ²Student COPUS codes: OG (group activity); AnQ (answering questions); SQ (student asks question); Ind (individual thinking); L (listening); WC (whole-class discussion); ³Instructor COPUS codes: MG (moving and guiding student work); AnQ (answering questions); PQ (posing questions); 1o1 (1-on-1 extended discussion with individual student); W (waiting); Lec (lecturing) (Smith et al., 2013).

3. Data and Results

3.1. Extent of Active Learning

Our dataset is comprised of COPUS observations of 64 classes, each taught by different instructors, in 35 academic departments at UHM. First, we calculated the PI to assess the degree to which active learning occurs in UHM classrooms. PI for the 64 classes ranged from 0-88%, with a mean of 43%, and standard deviation of 28%. These results compare favorably with the USA-wide results (mean=55%) reported by Stains *et al.*, (2018), which indicate that UHM courses are on average less didactic.

3.2. Correlations between Active Learning and Class Characteristics

No statistically significant differences were found in the PI of classes taught in STEM vs. non-STEM disciplines, or between Upper vs. Lower division courses (Table 2). Our key positive finding concerns class size: The difference in mean PI between large classes (PI=64%) and small classes (PI=39%) was highly significant ($p=0.02$; Table 2) and consistent with nationwide results (Stains *et al.*, 2018). In other words, instructors were statistically less likely to use active learning strategies in classes with >50 students.

Table 2. Results of Passivity Indicator (PI) analysis of 64 classes.

Class Characteristic ¹	PI Mean (stdev) ²	n ³	p-value ⁴
STEM (0)	49.1 (27.2)	47	0.23
Non-STEM(1)	27.7 (24.2)	17	
Upper (0)	43.4 (24.9)	34	0.53
Lower (1)	43.4 (31.4)	30	
Large (0)	64.1 (25.0)	11	0.02*
Small (1)	39.1 (26.7)	53	

¹STEM courses defined by NSF (2014); Upper and lower division courses defined as 300-400 and 100-200 levels respectively. Large classes have >50 students, small classes have <50 students (Freeman et al., 2014); ²PI mean (and standard deviation) shown as %; ³Number of classes; ⁴P-value, * indicates statistical significance at $\alpha=0.05$.

3.3. Enactment of CREDE Standards

As statistically measured by COPUS, the CREDE standards JPA and IC were only used during $\leq 6\%$ of class periods so were not further analyzed for this study. However, the COPUS codes for activities corresponding to enacting the CREDE standard LLD were used by students and instructors nearly a quarter of the time during small classes, and significantly less ($p=0.02$) during large classes (Table 3).

In small classes, students spent 24% of their time answering questions (AnQ) and instructors spent 20% of their time posing questions to students (PQ). In large classes, these figures were 12% and 11%, respectively. Similarly, in small classes, students only spent 50% of their time passively listening (L), which is not a CREDE activity, whereas in large classes they spent 66% of their time passively listening. All three differences between large and small classes were statistically significant, with p-values of 0.02-0.04 (Table 3).

Table 3. CREDE results for Language & Literacy Development as determined by COPUS.

LLD COPUS Codes ¹	SmalliClass n = 53 meani(stdev) ²	LargeiClass n = 11 meani(stdev) ²	p-value ³
(+) AnQ (s)	24.0 (15.1)	12.1 (11.3)	0.02*
(+) SQ (s)	7.3 (8.7)	3.6 (5.0)	0.18
(+) AnQ (i)	5.5 (6.9)	2.7 (3.9)	0.20
(+) PQ (i)	19.8 (11.6)	11.3 (8.7)	0.03*
(-) L (s)	50.4 (21.9)	66.2 (25.0)	0.04*
(-) Lec (i)	35.3 (20.9)	45.8 (21.1)	0.14

¹COPUS codes (see Table 1 for details; Smith et al., 2013) that match the Language & Literacy Development CREDE standard (CREDE, 2019); (+) indicates LLD is being enacted; (-) indicates LLD is not being enacted; (s) is for students, (i) is for instructors; ²Large classes have >50 students, small classes have <50 students (Freeman et al., 2014); n = number of classes; Mean % (and standard deviation) of class time spent doing the activity indicated by the COPUS code; ³P-value as calculated using student t-tests, * indicates statistical significance at $\alpha=0.05$.

4. Discussion and Recommendations

Despite compelling evidence that large classes can (*e.g.*, Deslauriers *et al.*, 2011) and should (Freeman *et al.*, 2014) be taught in an active way, active learning is not regularly occurring in large classes at UHM (this study) or elsewhere (Stains *et al.*, 2018). Our findings are alarming for two reasons: 1) Most students in our study were enrolled in large courses (n=1182 in 11 large classes, vs. n=1137 in 53 small classes); and 2) Active learning strategies have been shown to disproportionately benefit minority students (Haak *et al.*, 2011), and these students make up 34% of the UHM student body (US News and World Report, 2019). Although there is no specific recommendation for an optimal level of active learning, clearly a PI of 64% is too passive; large classes should strive for the PI reported for small classes.

Insight into the challenges that UHM faculty face in teaching large classes effectively comes from post-COPUS discussions with instructors of large courses, who revealed

sentiments of disbelief or anxiety about using active learning techniques: “*In large classes I struggle just to get through the material. Taking time to do group activities with 160 students means that I would get even farther behind.*”

For instructors accustomed to traditional lecture techniques, the prospect of completely revamping a high-enrollment course might seem daunting. Fortunately, small changes in instructional approach can yield significant, positive changes in student activity levels and learning outcomes, especially for diverse students (Haak *et al.*, 2011), with minimal effort from instructors (*e.g.*, Freeman *et al.*, 2014; Bruno *et al.*, 2017). Here is a sampling of easy-to-implement, active learning practices that have been shown to work effectively in large classes, with corresponding CREDE standards. References are provided for further review.

1. **Collaborative / 2 stage exams** (Gilley & Clarkston, 2014; Bruno *et al.*, 2017). Within a single class period, students take an exam twice: first on their own, and then in groups of 3-4 students that must agree on all answers. Enacts CREDE standard for JPA.
2. **Think/Pair/Share** (Lyman, 1981; SERC, 2019). Students THINK individually for a few moments about a question posed by the instructor; then PAIR up with another student to discuss their responses; then selected student pairs SHARE their ideas with the class. Enacts CREDE standards JPA, LLD, and IC.
3. **Group Worksheets** (Manjula *et al.*, 2010; CWSEI, 2013). Instructor creates a worksheet of questions to lead students through class content in a structured way. Make the first questions relatively easy, so that most groups know how to start, and make later parts more challenging. Enacts CREDE standard for JPA.
4. **Flipped Classrooms / Peer Instruction** (Crouch & Mazur, 2001; Smith *et al.*, 2009). Before class, students review course materials. In class, instructor poses questions based on pre-class preparation. Students answer individually. Instructor reviews student responses with whole class. In groups, students revisit question and discuss with their peers. Repeat until consensus is reached. Enacts CREDE standards JPA, LLD, and IC.

Faculty at UHM indicate a strong desire to teach less and instead conduct more research (ACCFSC, 2018). This presents an opportunity for instructors to turn their classrooms into test laboratories for evidence-based teaching techniques, with IRB approval. Two sections of a course taught by the same instructor could include a control section taught in a traditional lecture-only format, and an experimental section that included active learning or CREDE strategies. Student success measures could be compared between the two sections, possibly resulting in peer-reviewed publications to support progress towards tenure.

5. Conclusions

At UHM, we evaluated the use of active learning strategies across class types using COPUS, and found UHM classes to be less didactic (43% didactic) than nationwide averages (55%). However, large classes (>50 students) at UHM are taught significantly less actively than are small classes. This is also true with respect to the enactment of the CREDE standard LLD, which consists of question and answer sessions between instructors and students. Since active teaching strategies have been shown to work just as well in large vs. small classes, we present a range of activities that are easy to implement in large classes with minimal effort. We encourage faculty to use these techniques to better serve their diverse students, and also to create test laboratories within their classrooms to evaluate efficacy and disseminate results.

Acknowledgements

This project is supported by the US NSF's Research Infrastructure Improvement (RII) Track-1: 'Ike Wai: Securing Hawai'i's Water Future' Award #NSF/OIA-1557349; and by Improving Undergraduate STEM Education (IUSE) Award #NSF/GEO-1565950. The research protocol was approved as exempt the UH Institutional Review Board (#2017-003518). We gratefully acknowledge Daniel Port who assisted with statistical analyses.

References

- ACCFSC (2018). Faculty Worklife Survey. Retrieved from: www.hawaii.edu/offices/app/faculty/2018_Faculty_Worklife_Survey-Final.pdf
- Bruno, B.C., Engels, J., Ito, G., Gillis-Davis, J., Dulai, H., Carter, G., Fletcher, C. & Böttjer-Wilson, D. (2017). Two-stage exams: A powerful tool for reducing the achievement gap in undergraduate oceanography and geology classes. *Oceanography* 30(2), <https://doi.org/10.5670/oceanog.2017.241>.
- CREDE, (2019). Retrieved from: <https://manoa.hawaii.edu/coe/credenational/rubric/>
- Crouch, C.H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics* 69(9):970–977, <https://doi.org/10.1119/1.1374249>.
- CWSEI, (2013). In Class Activities tips. Retrieved from: http://www.cwsei.ubc.ca/resources/files/InClassActivities-tips_CWSEI.pdf
- Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science* 332(6031):862–864, <https://doi.org/10.1126/science.1201783>.
- Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H. & Wenderoth, M.P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the*

- United States of America 111(23):8,410–8,415, <https://doi.org/10.1073/pnas.1319030111>.
- Gilley, B.H., & Clarkston, B. (2014). Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching* 43(3):83–91.
- Haak, D.C., HilleRisLambers, J., Pitre, E. & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332:1,213–1,216, <https://doi.org/10.1126/science.1204820>.
- Jaccard, (1901). Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37, 241–272.
- Lorenzo, M., Crouch, C.H. & Mazur, E., (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics* 74(2):118–122, <https://doi.org/10.1119/1.2162549>.
- Lyman, F. (1981). The Responsive Classroom Discussion. In A. S. Anderson (Ed.), *Mainstreaming Digest* (pp. 109-113). College Park, MD: University of Maryland College of Education.
- Manjula D., Johnston, I.D., Johnston, H., Varvell, K., Robertson, G., Hopkins, A., Stewart, C., Cooper, I., & Thornton, R. (2010). Use of interactive lecture demonstrations: A ten year study. *Physical Review Special Topics - Physics Education Research* 6, 020119.
- NSF (2014). NSF Approved STEM Fields. Retrieved from: <https://www.btaa.org/docs/default-source/diversity/nsf-approved-fields-of-study.pdf>.
- Ruiz-Primo, M.A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L.A. (2011). Impact of undergraduate science course innovations on learning. *Science* 331:1,269–1,270, <https://doi.org/10.1126/science.1198976>.
- SERC (2019). Think-Pair-Share. Retrieved from: <https://serc.carleton.edu/introgeo/interactive/tpshare.html>
- Smith, M.K., Jones, F.H.M., Gilbert, S.L., Wieman, C.E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices, *CBE—Life Sciences Education* 12, 618–627.
- Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N. & Su, T.T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323:122–124, <https://doi.org/10.1126/science.1165919>.
- Springer, L., Stanne, M.E., & Donovan, S.S.. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology. *Review of Educational Research* 69(1):21–51.
- Stains, M., Harshman, J., Barker, M.K., Chasteen, S.V., Cole, R., DeChenne-Peters, S.E., Eagan Jr., M.K., Esson, J.M., Knight, J.K., Laski, F.A., Levis-Fitzgerald, M., Lee, C.J., Lo, S.M., McDonnell, L.M., McKay, T.A., Michelotti, N., Musgrove, A., Palmer, M.S., Plank, K.M., Rodela, T.M., Sanders, E.R., Schimpf, N.G., Schulte, P.M., Smith, M.K., Stetzer, M., Van Valkenburgh, B., Vinson, E., Weir, L.K., Wendel, P.J., Wheeler, L.B., & Young, A.M. (2018). Anatomy of STEM teaching in North American universities. *Science* 359(6383), 1468-1470.

- Tharp, R. G., Estrada, P., Dalton, S. S., & Yamauchi, L. A. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder, CO: Westview.
- US News & World Report, (2019). *Campus Ethnic Diversity*. Retrieved from: www.usnews.com/best-colleges/rankings/national-universities/campus-ethnic-diversity
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Yamauchi, Lois A.; Taira, Kazufumi; Trevorrow, Tracy (2016). Effective Instruction for Engaging Culturally Diverse Students in Higher Education. *International Journal of Teaching and Learning in Higher Education*, v28 n3 p460-470.