

September 30, 2019

To: Henning Bohn, Chair  
Academic Senate

From: William Smith, Chair, ESCI *ad hoc* Committee , MCDB  
Linda Adler-Kassner, Associate Dean, Undergraduate Education, Writing Program  
Stanley M. Awramik, Earth Science  
Maria Charles, Sociology  
Rachael King, English  
Everett Lipman, Physics  
Jeff Moehlis, Mechanical Engineering  
Vickie Scott, Theater and Dance  
Rene Weber, Communication

Re: Report of the ESCI *ad hoc* Committee

The *ad hoc* Committee on Student Evaluation of Teaching was established Fall Quarter 2018 in response to a memo from the Undergraduate Council (Appendix 1). The Council raised concerns about the effectiveness of our current Evaluation System for Courses and Instruction (ESCI) in evaluating instructional quality, its pedagogical consequences, and its possible negative bias toward women, racial and ethnic minorities, and other groups in the results. Accordingly, the committee first tasked itself with thoroughly assessing the current system. In doing this the committee employed a number of sources including peer-reviewed studies published in journals in higher education and other disciplines (see bibliography). and consultation with UCSB faculty and staff who administer the ESCIs and participate in course improvement. We also consulted with key participants in UCSB's faculty merit and promotion review process, and with representatives of the UC Santa Barbara Faculty Association. The second task of the committee was to research systems at other universities, commonly referred to as Student Evaluation of Teaching (SET), both within and outside of the UC system. The committee met nine times last academic year, either as a whole or as one of two subcommittees. Though the committee's work is still in progress, we offer the following report.

## History of UCSB's Student Evaluation of Teaching System

The current ESCI system dates back to the 1970s and was originally aimed at providing feedback to instructors for course improvement. In the following decades, the ESCIs were incorporated into the merit and promotion process. It is clear to us that ESCI survey results are not sufficiently robust and meaningful to fill such a role.

Academic personnel policy, as detailed in the Red Binder, explicitly states what is to be used in personnel reviews. For Senate faculty, the Red Binder sections on Routine and Expanded Merit and Promotion Cases ([I-31.V.](#) and [I-34.VI.](#)) state the following:

“Evaluation of the teaching record. At a minimum, two sources must be included in the case. ESCI summary sheets and scores for questions A and B are mandatory<sup>[1]</sup>.”

For Unit-18 Lecturers, the MOU between UC and the American Federation of Teachers [Article 7b](#), entitled “Process for Achieving Continuing Status and Continuing Appointments,” also outlines what is to be included in the reviews of Non-Senate faculty (NSF). Article 7b states that:

“Evaluations of the academic qualifications or performance of NSF for purposes of consideration for a continuing appointment shall be made on the basis of demonstrated excellence in the field and in teaching, academic responsibility, and other assigned duties which may include University co-curricular and community service.”

The MOU also states that:

1. Instructional performance is measured by evaluation of evidence demonstrating such qualities as:
  - a. command of the subject matter and continued growth in mastering new topics;
  - b. ability to organize and present course materials;
  - c. ability to awaken in students an awareness of the importance of the subject matter;
  - d. ability to arouse curiosity in beginning students and to stimulate advanced students to do creative work

For excellence (“continuing”) reviews, the MOU provides additional explicit language:

2. All relevant materials shall be given due consideration. These may include:

- a. student evaluations, *provided that the quantitative measure in the student evaluation is not the sole criterion for evaluating teaching excellence*;
  - b. assessment by former students who have achieved notable professional success;
  - c. assessments by other members of the department, program or unit, and other appropriate faculty members;
  - d. development of new and effective techniques of instruction and instructional materials; and
  - e. assessments resulting from classroom visitations by colleagues and evaluators.
3. A NSF may provide a self-statement or self-evaluation of her/his teaching objectives and performance.

(Italics added)

Cognizant that the ESCI system is deeply ingrained in our current academic personnel system for Senate and NSF alike, the committee acknowledges that any changes will need to be planned and implemented carefully and in consultation with all relevant parties. We hope that the findings and recommendations presented in this report will help inform this process.

### **Ad Hoc Committee Evaluation and Findings**

Much of the committee's initial work was accomplished through its two subcommittees. The Committee's findings are discussed below:

#### **ESCI Validity, Reliability, and Use**

- *Published peer-reviewed studies have shown that evaluation systems similar to UCSB's do not give valid, reliable and objective measures of teaching effectiveness.*

All Senate faculty and NSF continuing lecturers undergo regular evaluation for merit increases. Ladder faculty also undergo career reviews for promotion to Associate and Full professor (and beyond); Senate lecturers undergo career reviews from L(P)SOE to LSOE to SLSOE. A key component of these cases, regardless of faculty status, is the documentation of teaching effectiveness, and numeric ESCI scores have historically been relied upon as a central measure of teaching quality. However, a growing body of literature has documented that surveys asking students to rank the quality of instruction on a numerical scale do not provide an accurate measure of teaching effectiveness (eg., Braga and Paccagnella 2014, Stark and Freishtat 2014, Boring, Ottoboni et al. 2016, Rivera and Tilcsik 2019). For example, Braga and Paccagnella

examined the correlation between SET scores and how well students were prepared in one course for the second course in two-course series. Surprisingly, an inverse relationship was found between high SET scores and subsequent academic performance. Comparisons in introductory physics classes of SET with objective measures of student learning have shown an inverse relationship and no relationship, respectively, between the two quantities (Lee et al., 2018). This is consistent with a limited study of introductory physics classes at UCSB (see Appendix 2). Although we do not have the means to presently undertake a comprehensive study here at UCSB, the weight of published data from comparable universities (e.g., Stark and Freishtat 2014 on UC Berkeley) points to a significant problem with our current merit and promotion system that relies so heavily on ESCI scores.

Additionally, because the system generates a numerical score, it gives a false sense of precision. As a measurement, ESCI scores are uncalibrated, and, at best, accurately reflect only the extremes of instructional quality.

- *Published/peer-reviewed studies have shown that evaluation systems similar to UCSB's have an inherent bias against certain groups of instructors.*

Research indicates that SET scores can be negatively biased by such factors as the time of day the course is taught, course subject matter, and, most troublingly, the race, gender, or sexual orientation of the instructor (Miles and House 2015, Mengel, Sauermann et al. 2018, Mitchell and Martin 2018, Rivera and Tilcsik 2019). For example, one study found that the language “students use in evaluations regarding male professors is significantly different than language used in evaluating female professors” (Mitchell and Martin 2018). Another experimental study found that students rated online instruction lower when they were told the instructor was female than when they were told the instructor was male (MacNeill 2014). Students also tend to evaluate professors teaching quantitative courses (e.g. math or statistics) less favorably (Uttl and Smibert 2017). The American Sociological Association, with the endorsement of 19 other professional societies, has issued a set of recommendations based on the extant research attesting to the biases reflected in conventional SETs (ASA 2019).

- *Students have used the anonymity of the comments section of the ESCI form to make irrelevant, sometimes inappropriate, sexist, or racist comments.*

While our data on this problem at UCSB is anecdotal, this issue has been raised in a number of articles (e.g., Schmidt 2017, Falkoff 2018), suggesting that it is widespread. While we must respect the right of students to comment freely, the committee members agree that better education of incoming students on how the evaluation data is used, and on how to comment constructively may at least partially alleviate this problem (see Recommendations below).

- *Our current system does not support faculty members' instructional goals; rather, it rewards popularity more than rigor and incentivizes instructor “likability,” sometimes at*

*the expense of learning. This is especially true for topics that are not liked by students or perceived as difficult.*

Any assessment system, whether within a course or of teaching, shapes the instruction within that course. UCSB faculty seek to create courses that are rigorous, challenging students to grow as learners and thinkers. The current ESCI questions, however, reward “popularity.” Alternative models, for instance, ask students about their understanding of course goals, the extent to which instructional activities support those goals, and what contributed to their intellectual growth. Because the currently-mandated ESCI questions provide minimal guidance with regard to what is actually being rated, they tend to incentivize making students happy. This could contribute to phenomena such as grade inflation, lowering of standards, and/or dilution of curricula. One study, for instance, showed that passing out cookies in class significantly increased SET scores (Hessler et al. 2018). Institutional pressures on structurally vulnerable members of our teaching faculty to deliver high SET scores may also contribute to broad changes in the learning environment and campus culture and may negatively affect learning outcomes.

- *While we should value student opinions, surveys should gather information from students about their experience in the course rather than about the quality of teaching.*

Surveys on student experiences in a course can contribute to faculty members' understandings of how students perceive a course's instructional goals and the practices employed by faculty to achieve those goals. Faculty can then take this input into account as they seek to assess and improve their own teaching. For example, a first-year student fresh out of high school does not have the perspective necessary to evaluate the content of a university-level course (ESCI question "B"). We would not measure the success of a surgery by surveying the patient in the recovery room. Likewise, our teaching evaluation system should recognize the limitations of student opinion.

In summary, the committee has unanimously reached the conclusion that our current system for SET is neither valid nor reliable, is biased, and negatively influences the campus learning environment. ESCIs were not designed to serve the purposes for which they are currently being used, and their meaning is widely misunderstood by those who provide and use them. This lack of understanding informs every aspect of the ESCI/feedback process, from survey construction, to student response, to use in the academic personnel process.

SET can be used for a variety of purposes: to provide formative feedback to an instructor that is intended to inform and improve their teaching; to provide insight into students' experiences in a course; or to provide summative feedback about students' perceptions of course effectiveness. Since guidance for students and faculty regarding the purpose(s) of ESCIs and how the formulation of survey items might be used is minimal, there is little evidence to attest to any shared sense of understanding of those purposes. A foundational understanding, shared among all involved in data gathering and use, is essential for both surveys and results to be valid.

Furthermore, the duality of purpose for the ESCIs (course improvement and instructor evaluation) raises questions as to whether the ESCI system is meeting its intended outcomes – in either role.

## **Recommendations for an Improved Student Evaluation of Teaching System**

### **Considerations**

In the next sections, the committee presents recommendations for addressing what we unanimously agreed to be unacceptable deficiencies in our current system. These recommendations parallel many included in the American Sociological Association report, which is itself based on a synthesis of existing research and practice. The following considerations were in the foreground of the committee's discussions. Recommended changes must:

- Take into consideration language in the Red Binder for both Senate and NSF faculty that emphasize the use of *at least two forms of evidence to attest to instructional effectiveness*. While revision of mandated ESCI questions may take time, the campus can create guidelines that more explicitly outline what “second forms” may take, and how reviewing agencies can evaluate these “second forms” in valid and reliable ways.
- Consider and adopt, where appropriate, the results of parallel studies by other UC agencies. The Instructional/Faculty Development professionals have completed a thorough review of the relevant policy and literature and are in the process of making recommendations. In addition, the systemwide Senate is set to take up questions associated with SETs in 2019-20.
- Be feasible and scalable to a large university like UCSB.
- Take into account the dual contribution of student evaluations, (*i.e.*, course improvement and instructional evaluation).
- Recognize the role of student input in instructional improvement. Students should be informed that the faculty take their feedback into consideration.
- Be based on evidence, use scientific survey methodology, and ensure that SETs reflect principles of validity from the research literature. Reform is essential, but we need to be confident that changes to the current system will address the concerns outlined above. Thus, we should resist the temptation of “change for the sake of change.”
- Not lead to undue increase in faculty, staff, or student workload.

- Be made with sufficient notice, so as to protect Senate and non-Senate faculty who will be undergoing review in the next several years and who have built their cases with the current system in mind.

### **Short-term Changes**

Even if it is widely agreed that changes to our ESCI system are necessary, planning and implementation could take several years. Nevertheless, the problems this committee uncovered with UCSB's current system are so serious that we feel immediate action is imperative. In fact, some committee members are proponents of a moratorium on the use of ESCI scores for merit and promotion cases until an improved system is in place. However, because of the disruption this could cause to current and near-term cases, it may be that qualified use of ESCI's scores will need to continue in the interim. Our short-term recommendations below, represent what the committee feels is the minimum action necessary to help mitigate these problems:

- 1) Department chairs, deans, and other campus reviewing agencies be provided with this memo as soon as possible, and be encouraged to consult with their faculty regarding alternative or additional documentation of teaching effectiveness (this is consistent with Red Binder language about two forms of documentation). Instructional Development and/or the Center for Innovative Teaching, Research, and Learning are available to consult about these possibilities.
- 2) Reviewing agencies and administrators should take seriously the guidelines for interpreting ESCI scores provided by Instructional Development and be aware of the limitations of the evaluations. For instance, departments and other reviewing agencies should be cautioned against assigning meaning to variations in mean scores unless they are extreme.
- 3) Cease comparisons of individual instructor scores with campus averages and department averages that include dissimilar courses.
- 4) The campus should provide more structured guidance on the importance of teaching evaluations and how to provide effective and appropriate feedback. For example, UCSB might look to models such as those developed by UC Merced, which uses videos created by and featuring undergraduate students about how to provide feedback on teaching: [https://cetl.ucmerced.edu/SATAL\\_Video](https://cetl.ucmerced.edu/SATAL_Video). McGill provides a similar video which centers on implicit bias: <https://mcgill.ca/mercury/about/equity>. A similar video targeted to UCSB students, along with changes in survey design, could help to improve the quality of feedback provided to faculty via SETs.

## Road Map for Substantive Changes to the ESCI System

In addition to the short-term actions outlined above, longer-term action is necessary:

- 1) *For Senate and NSF merit and promotion cases, develop guidelines and definitions for excellence in teaching.*

Before the campus can consider revisions to the ESCI system, the relevant stakeholders need to agree on what we are attempting to measure. The current ESCI questions A and B are often interpreted as a measure of student satisfaction and student happiness. While satisfaction can remain as one element of instructional assessment -- keeping in mind that “satisfaction” and “challenge” are simultaneously possible -- the committee agrees that faculty, students, and stakeholders must define key goals, attributes, and characteristics associated with “learning.” Once these are developed, they can then be used by departments and campus reviewing agencies in interpreting survey results.

- 2) *Revise ESCI/SET questions based on new guidelines and definitions for excellence in teaching.*

Our current SET system, the ESCI, asks students to rank both the instructor (Question A) and then the course independent of the instructor (Question B), on a 1-5 scale. Departments are permitted and encouraged to include additional questions, although the use of these is not required by the Red Binder for merit and promotion cases. As stated above, this system has remained largely unchanged for several decades. This committee believes that it is essential to reassess and revise our student surveys once we have agreed-upon parameters of instruction that we are seeking to measure.

The *ad hoc* committee has researched how other campuses have revised their SETs. We find that recent revisions made to student surveys at UC Santa Cruz are a possible example of the direction we could recommend (Appendix 3). The UCSC survey first gathers information on the student (without personal identifiers) regarding why they are taking the particular course and how much effort they have put into the course. The survey then presents a series of statements such as, “The instructor communicated and explained concepts clearly.” The student can then choose from these responses: “unable to comment/never/occasionally/somewhat frequently/frequently/very frequently.” The goal is to measure what the student got out of the course, rather than simply ranking the instructor on an excessively granular scale. The committee supports revising our survey along these lines and, importantly, using students’ motivations and effort variables as statistical controls (i.e. reporting scores controlled by student motivation and effort).

Additionally, a recent published peer-reviewed study from Iowa State University reported that changes to the wording of student surveys can significantly reduce gender bias (Peterson et al., 2019). Another suggests that seemingly minor technical aspects of survey design such as



changing the length of the rating scale (e.g., from a scale of 10 to a scale of 6) significantly reduced the amount of gender bias (Rivera and Tilcsik 2019).

This recommendation raises the question of who in particular will draft the revised questions. Because of the importance of this project, the committee recommends a commitment of resources. The leaders of this ongoing effort must be Academic Senate members with knowledge and expertise in higher education theory and practice, and in scientific survey methods. We recommend the establishment of a two-year research project staffed with 1-2 senior experts in this research area at the professor level. This project will require funds for both teaching release and for staff support. Any revisions should be made with input from faculty, non-senate instructors (or their representatives) and students. Prior to deployment of any evaluation method or new survey questions, there should be independent scientific validation by someone other than the developer.

*3) Consider numerical scores only in the context of other measures of teaching excellence (which should replace or supplement the two measures currently required by the Red Binder).*

Although the M&P case material guidelines in the Red Binder require that at least two sources be submitted for evaluation of teaching, one of them must be the ESCI scores. Other forms of allowable documentation are discussed in Red Binder [I-75-V-2](#) for Senate faculty and Red Binder [II-6](#) and [II-10](#) for non-Senate faculty.

Due to the well-documented problems with numerical scoring of instructors in the scientific literature, we recommend elimination, or at a minimum deemphasis of ESCI scores in merit and promotion cases. After further research and evidence (see below), ESCI data should only be one of several sources used for evaluating teaching. This would allow faculty who feel that ESCI data would not give an objective measure of their teaching effectiveness due to bias, or other factors such as class size or type, to deemphasize or omit these scores without stigma.

*4) Periodically Re-Evaluate SETs and Teaching Evaluation Processes*

One of the biggest challenges we face in making recommendations for revising UCSB's SET system is that while there is extensive documentation on the shortcomings of numerical scoring of instruction by students, there are few studies that document alternatives as being better – either in providing meaningful data on which to base instructor advancement or in reducing bias. Because many institutions have only recently made changes to their SET systems (the University of Oregon being a particularly visible example (Doerer 2019), it may be a number of years before we know whether these changes have had the desired effect. The committee feels strongly that periodic reevaluation is necessary if changes are made to UCSB's SET system. However, in order for periodic reevaluation to be meaningful, it will be necessary to establish a baseline from which improvement can be measured.

5) *UCSB should move to an all-electronic ESCI system.*

There is currently an ESCI Online Pilot Project at UCSB. Online course evaluations are now being used in a number of UCSB courses and are routine for teaching assistant evaluations. Pending the final report of the task force assessing the pilot project, the committee believes that moving to an all online system is an inevitability, and will put UCSB in line with other institutions. We anticipate that specific recommendations on how to best implement an online system will be made by the task force.

The collection of SET data electronically raises the possibility of less disruption of courses, increased student input, and more advanced analyses of the data. For example, collecting student written comments digitally could have the benefit that responses for use in merit and promotion cases could be appropriately filtered, analyzed, and more objectively summarized with state-of-the-art natural language processing (NLP) technology. For example, McGill University has established a protocol for course evaluations that are “determined to be hateful or discriminatory,” which includes an option for the instructor to request that the student’s response be struck in its entirety from the course evaluation.

### **Conclusion**

The *ad hoc* Committee on Student Evaluation of Teaching through its work in the past academic year has reached the conclusion that serious deficiencies are present in UCSB’s evaluation system, and that action is urgently needed. The changes recommended by this committee, both short-term and long-term, will help support UCSB’s continued excellence in undergraduate instruction, and will promote the dual goals of diversity and fairness in instructor evaluation and promotion.

## Bibliography

American Sociological Association. (2019). "Statement on Student Evaluations of Teaching."

Boring, A., et al. (2016). "Student Evaluations of Teaching Mostly DO Not Measure Teaching Effectiveness." ScienceOpen.

Braga, M. and M. Paccagnella (2014). "Evaluating Students' Evaluations of Teachers." Economics of Education Review: 71-88.

Doerer, K. (2019). "Colleges Are Getting Smarter About Student Evaluations. Here's How." The Chronicle of Higher Education: January 13, 2019.

Falkoff, M. (2018). "Why We Must Stop Relying on Student Ratings of Teaching." The Chronicle of Higher Education.

Flaherty, C. (2018). "Teaching Eval Shake-Up." Inside Higher Ed.

Hessler, M., et al (2018). "Availability of cookies during an academic course session affects evaluation of teaching." Medical Education 52 (10): 1064-1072.

Lee LJ, Connolly ME, Dancy MH, Henderson CR, Christensen WM (2018). "A comparison of student evaluations of instruction vs. students' conceptual learning gains." The American Journal of Physics 86(7), 531-535.

MacNeill, L. (2014). "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching." Innovative Higher Education 40(4): 291-303.

Mengel, F., et al. (2018). "Gender Bias in Teaching Evaluations." Journal of the European Economic Association 17(2): 535-566.

Miles, P. and D. House (2015). "The Tail Wagging the Dog; An Overdue Examination of Student Teaching Evaluations." International Journal of Higher Education 4(2): 116.

Mitchell, K. and J. Martin (2018). "Gender Bias in Student Evaluations." PS: Political Science & Politics 51(3): 648-652.

Owen, Ann. "The Next Lawsuit to Hit Higher Education."  
<https://www.insidehighered.com/views/2019/06/24/relying-often-biased-student-evaluations-assess-faculty-could-lead-lawsuits-opinion>

Peterson DAM, Biederman LA, Andersen D, Ditonto TM, Roe K (2019) Mitigating gender

bias in student evaluations of teaching. PLoS ONE 14(5): e0216241.

Rivera, L., & Tilcsik, A. (2019). "Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation." American Sociological Review 84:248–274.

Schmidt, P. (2017). "When Students' Prejudices Taint Reviews of Instructors." The Chronicle of Higher Education.

Stark, P. and R. Freishtat (2014). "An evaluation of course evaluations." ScienceOpen Research.

Uttl, B. and D. Smibert (2017). "Student evaluations of teaching: teaching quantitative courses can be hazardous to one's career." PeerJ — the Journal of Life and Environmental Sciences 5:e3299.

McGill University. Equity and Course Evaluations. (n.d.) Retrieved from:  
<https://mcgill.ca/mercury/about/equity>

---

[1]“Q1: In rating a course, you should consider the instructor’s teaching apart from the course materials or content. Please rate the overall quality of the instructor’s teaching.“ Q2: Please rate the overall quality of the course, including its material or content, independent of the instructor’s teaching.”

**Academic Senate  
Santa Barbara Division**

May 22, 2018

To: Henning Bohn, Divisional Chair  
Members of the Executive Council

From: David Paul, Chair  
Undergraduate Council

Re: Evaluation System for Courses and Instruction

Undergraduate Council suggests the formation of an ad hoc committee or taskforce to look more closely at the role of ESCI evaluations. We imagine that such a group would make an assessment of whether Academic Senate should continue to support ESCIs in their current form, revise them, or possibly replace them with other modes of teaching assessment.

Undergraduate Council is concerned about the extent to which the campus has become reliant on the ESCI evaluations for gauging teaching performance, a development that has ramifications for both the quality of instruction and the promotion/merit-review process at UCSB. We have three specific concerns: the statistical limitations of the ESCIs; misunderstandings about what they measure; and the body of empirical evidence demonstrating that student evaluations like the ESCI display biases against faculty of color, female faculty, and faculty from other underrepresented groups (for review, Boring and Stark, 2016). We worry that instructors (particularly those whose employment hinges primarily or exclusively on their teaching performance) alter classroom practices to game ESCI results, sometimes to the detriment of pedagogical outcomes.

Instructional Development, which oversees the ESCI process, has been quite clear about the statistical limitations of the evaluations. But some departments proceed without regard for these limitations, citing numerical results of the survey even in instances where the class size is so small as to render them meaningless. In departments where some instructors teach exclusively small classes, the overall departmental average is skewed, potentially with negative impacts on faculty in the same department teaching large classes, where the anonymity of numbers makes students feel more comfortable providing a negative ESCI evaluation.

Second, although prior research on students evaluations of teaching (SET) ratings suggested that they are a modest predictor of teaching effectiveness, a recent meta-analysis (Uttl, White & Gonzalez, 2016) of numerous studies investigating this issue found that prior conclusions were likely a product of inadequately powered studies and publication bias. With a more exhaustive evaluation of the evidence, there was no significant correlation between SET ratings and learning. The authors of this important study concluded, "These findings suggest that institutions focused on student learning and career success may want to abandon SET ratings as a measure of faculty's teaching

effectiveness.” A number of other studies have also found SETs problematic (Stark & Freishatat, 2014; Hornstein, 2017). In short, what the ESCI actually measures is the sense students have of their classroom experience, not the quality of teaching. Thus ESCI evaluations are affected by students’ perceptions of an instructor’s general disposition or whether the instructor is an easy grader, neither of which is necessarily an indicator of the quality of teaching. As a result, instructors’ incentives are skewed toward making students happy instead of towards focusing on the teaching that would be most beneficial to students.

Third, and most disturbingly, a number of studies have demonstrated that quantitative course evaluations have problems of bias with respect to faculty of color, female faculty, and faculty of other underrepresented groups (Boring, Ottoboni, & Stark, 2016). We feel that relying so heavily on as imperfect a diagnostic tool as the ESCIs for teaching evaluations works to the detriment of the University’s goal of fostering a vibrant faculty that reflects the diversity of California.

#### References:

Boring, Anne, Kellie Ottoboni, and Philip B. Stark. “Student Evaluations of Teaching (mostly) do not Measure Teaching Effectiveness.” *ScienceOpen Research* 2016.

<https://www.scienceopen.com/document/read?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e>

Hornstein, Henry A. “Student Evaluations of Teaching are an Inadequate Assessment Tool for Evaluating Faculty Performance.” *Cogent Education* 4, no. 1 (2017).

<https://www.tandfonline.com/doi/abs/10.1080/2331186X.2017.1304016>

Stark, Philp B. and Richard Freishtat. “An Evaluation of Course Evaluations.” *ScienceOpen Research* 2014. <https://www.scienceopen.com/document/read?vid=42e6aae5-246b-4900-8015-dc99b467b6e4>

Uttl, Bob, Carmela A. White, and Daniela Wong Gonzalez. “Meta-analysis of Faculty’s Teaching Effectiveness: Student Evaluation of Teaching Ratings and Student Learning are not Related.” *Studies in Educational Evaluation* 54 (2017): 22-42.

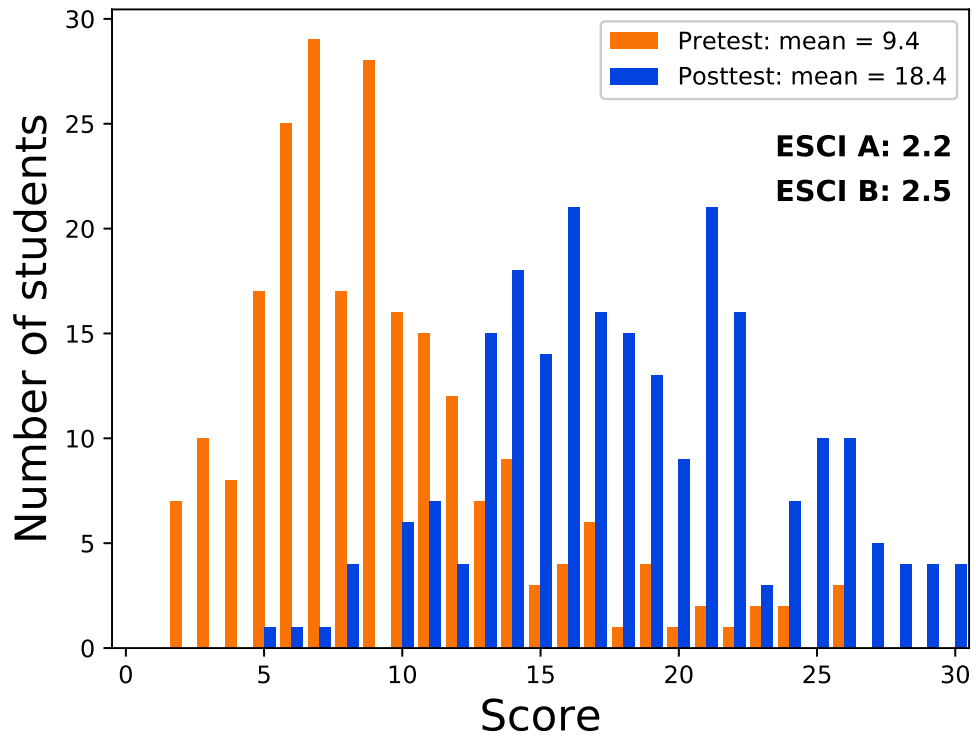
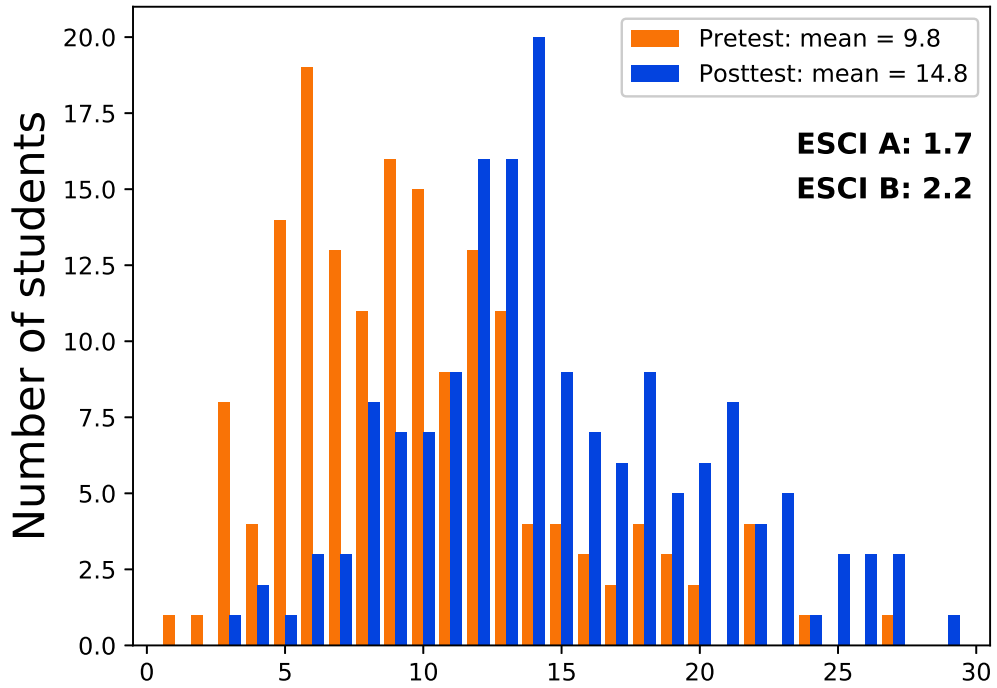
<https://www.sciencedirect.com/science/article/pii/S0191491X16300323>

CC. Debra Blake, Executive Director  
Shasta Delp, Assistant Director

### **Introductory Physics Course Study**

In the Fall quarter of 2003, identical pre- and post-tests were given to two sections of UCSB's Physics 6A, an introductory course for life sciences majors. The test was the Force Concept Inventory, a nationally recognized set of questions designed to probe student understanding of Newtonian mechanics. The accompanying histograms illustrate how the class whose instructor got ESCI A and B scores of 2.2 and 2.5, respectively, learned significantly more than the class whose instructor got scores of 1.7 and 2.2.

# UCSB Physics 6A: Assessment vs. ESCI





**February 7, 2019**

*Note: This version represents the collaboration between COT, CAAD, CAP, IRAPS and Jody Greene following the 2.5.19 SEC meeting.*

**STUDENT INFORMATION**

1. What is your current class standing at UCSC?

- Freshman/first year
- Sophomore/second year
- Junior/third year
- Senior/fourth year
- Fifth-year senior or more
- Master's student
- PhD student
- Other

2. Why are you taking this class?

- Required for my major/minor
- Elective for my major/minor
- Part of a proposed major/minor I am exploring
- To fulfill a GE requirement (outside my major/minor)
- General interest in the topic
- Other reasons

3. What percentage of class meetings taught by this instructor (in person or remotely, not counting sections or labs taught by others) did you attend? (Note: 1 week = 10%)

- I withdrew from the course.
- 0-24%
- 25-49%
- 50-74%
- 75-100%

4. About how many total hours per week, outside of class meetings, did you spend on work for this course?

- 0-3 hours
- 4-6 hours
- 7-9 hours
- 10-12 hours
- 13 hours or more

**The purpose of this anonymous survey is:**

1. To give you a chance to reflect on how your experience with your instructor influenced your learning in the course;
2. To give your instructor feedback that may be helpful in improving the effectiveness of their **instruction** or the **design** of this course.
3. To give university administration and instructor's department/program/college evidence of your instructor's teaching effectiveness for their personnel reviews.

The instructor will not see responses until after grades have been submitted.

Please **only comment on your experience with the primary instructor**. Please fill out a **separate survey for any teaching assistants** for this course.

**FEEDBACK ON INSTRUCTION:**

(Scale for 5-9 is: unable to comment/never/occasionally/somewhat frequently/frequently/very frequently)

5. The instructor used class time effectively to support my learning.
6. The instructor communicated and explained concepts clearly.
7. The instructor showed respect for all students in the class.
8. The instructor provided useful feedback on my assigned work (put "unable to comment" if you received feedback on your assignments only from a Teaching Assistant).
9. The instructor clearly communicated how assignments would be evaluated and/or graded.

**FEEDBACK ON COURSE:**

(Scale: never understood the goals/at the beginning of the course/at the end of the course)

10. I understood the learning goals or learning objectives of the course.

(Scale for proposed 11-13 is: unable to comment/never/occasionally/somewhat frequently/frequently/very frequently)

Please only comment if the course contained the specific activity addressed in questions 11-13. Otherwise select “unable to comment.”

11. Lectures and other instructor-led presentations were well structured and had clear goals.

12. In-class activities were well structured and had clear goals.

13. Problem sets, writing assignments, and other homework helped me feel prepared for examinations, papers, and projects.

(New scale for question 14 is: no assigned reading/I did little to none of the assigned reading/I found the reading somewhat useful/I found the reading useful/I found the reading very useful)

14. I found the assigned reading I completed to be useful to my learning in the course.

#### **Comments OPEN-ENDED**

15. Please describe any specific **teaching practices** (such as lectures, seminar discussions, small group activities, individual conferences) the instructor used that

16. Please describe the specific **course elements** (including readings, films, homework, guest lectures, instructional videos, examinations, papers, study guides, or other elements) that were helpful or unhelpful to your learning in this course.

17. What suggestions, if any, do you have to improve this course? Please be as specific as possible.

18. Is there anything else you would like to add?

#### **Preparation for the Course**

19. Did you feel prepared, by prerequisites or prior coursework, for the work required in this course?

- Unable to comment
- Not at all prepared
- Somewhat prepared
- Prepared
- Very prepared

#### **Comments OPEN-ENDED**

20. Please restate your answer to Q19 and explain it. For example, I felt somewhat prepared because ....