Student Evaluations in the Information Studies Department: Issues and Recommendations

IS 289-5
Rebecca Fordon
Dominique O’Neill
March 14, 2016
Executive Summary

Student evaluations have long been a way to provide students a voice in their education, bridging the gap between student experiences and the instructor’s teaching methods. However, since their origins as a tool for teaching, they have acquired additional uses as new stakeholders have taken advantage of the information they provide. The university administration uses them to inform employment decisions. Students might like to use them to inform curriculum choices. And of course, faculty do still use them to receive feedback about teaching effectiveness. Despite these multiple purposes, the design of the system for collecting and disseminating formal student evaluation in UCLA’s Information Studies is largely oriented towards the administration’s purpose of informing employment decisions.

Objective: This report will address student evaluations within the UCLA Information Studies department. Specifically, we will address the technical and social contexts in which student evaluation information is produced, transmitted and used. We will also discuss the governance of student evaluation information, including the laws and political processes that shape how the information is handled. Finally, we will examine trends and points of resistance that may inform how student evaluation information may be used in the future.

Process: We researched the context of student evaluations in the IS department by conducting interviews with many of the people involved in producing and handling the evaluations, such as IS faculty, the chair of the IS department, staff of the Office of Instructional Development (which is charged with overseeing student evaluations at UCLA), and the staff that help compile student evaluation information for transmission to faculty and for use in employment decisions. We also conducted a survey of students in Data Informatics, in which we asked questions about the students’ interest in access to student evaluation information. Finally we researched how other institutions handle student evaluation information.

Main recommendations: Our primary recommendations relate to how IS faculty and the IS department chair can better use the data contained in student evaluations. Currently, the department chair receives summary information only in PDF form; we recommend that it be provided in .csv format. We also recommend some visualization tools and techniques that both faculty and the department chair should find useful. In the long-term, we also recommend that the IS department consider different questions that could provide a better view into teaching effectiveness.

As for student access to evaluations, we recommend that rather than trying to use student evaluations (a tool designed to give feedback regarding teaching effectiveness) to choose courses (a calculus that involves far more than simply teaching effectiveness), that students collaborate with the department on other tools that could inform student choice. These tools could include some of the options that other teams are considering (such as making course syllabi more
available), social collaboration that is specifically designed to minimize the introduction of bias, as well as other tools that we discuss further in our report.

**Introduction to Student Evaluations at UCLA**

Student evaluations have a long history at UCLA, dating back to the 1970s. Their original purpose was to provide students a voice and to increase teaching effectiveness. UCLA’s Office of Instructional Development (OID), through its Evaluation and Assessment Program (EAP) has provided student evaluation services to departments at UCLA for at least the past 30 years. For many years, these evaluations were conducted by paper, consisting of a sheet handed out towards the end of the quarter that would ask the student to rate the class in various categories (quantitative data) as well as to provide written comments (qualitative data). OID would calculate averages of the ratings and would provide those, along with the completed paper forms, to the departments for distribution to instructors.

Today, most student evaluations are conducted online. Students either receive a link to a questionnaire and verify their identity or login into their MyUCLA account, and then they answer the same sorts of quantitative and qualitative questions as were required for paper evaluations. OID stores the answers in a database, and provides summaries to a department liaison, who distributes them to the department chair and instructors.

In Information Studies (as in most departments), students do not have access to any of the data in student evaluations. Students therefore sometimes turn to other sources of information to make decisions about course selection. These may include third-party websites (such as ratemyprofessors.com and bruinwalk.com) as well as social media and in-person conversations.

**Context**

This section will summarize the information and circumstances surrounding student evaluations in the Information Studies department, including the production, standardization, storage, access, circulation, visualization, long-term preservation of student evaluations.

**Production**

The production and completion of student evaluations is managed by OID and produced by students. The Information Studies department uses a specialized form (Form E) consisting of specialized evaluation questions (Form E) originally developed by the Education Department. Evaluation questionnaires are disseminated to students through their MyUCLA account and email. Once they login, they have the choice to fill out and submit the form. Completing the forms is at the discretion of the students and is not required. However, this can lead to low completion rates; some estimates of current completion rates are as low as

---

1 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
While the switch from paper to online evaluations has decreased the completion rate, it has also decreased the amount of labor that goes into production. Estimates of the number of paper evaluations, which in the previous system had to be inputted and collated, were 300,000 sheets of paper and 18,000 envelopes for undergraduates alone. OID was responsible for physically collecting these documents, inputting the information into computers, compiling it into reports and disseminating it on CDs to the evaluation coordinator for each department. Now with the switch to online evaluations, the information gets directly deposited into OID’s database where they collate the data into reports. They produce three different reports: rundown reports, department summary reports, and instructor summary reports. They also include a folder with data files for each instructor. The rundown reports are in PDF format and list instructor averages by course. Department summary reports are also in pdf format and list each question with the averages, median, and standard deviation and number of student responses for each question contained in the evaluation averaged over the entire department. The final reports distributed are individual instructor reports that come in both PDF and CSV format. These produced files then get uploaded to each department coordinator’s box.com account.

**Standardization**

The primary form of standardization for student evaluations comes in the implementation of a three-point rating scale for all questions, and in the standard selection of questions across the whole department. There is no ability for instructors to add custom questions.

One of the biggest problems with this standardization is that it lacks granularity—it is difficult to distinguish (for example) between a 2.75 and a 2.80. Most professors interviewed cited this lack of granularity as the reason why they do not find these reports particularly useful for informing their teaching methods. There is no way to calibrate what these averages signify, they fluctuate from professor to professor. Also in relation to each other they are so similar that there is no way to calibrate areas that need improvement. For instance in the Department Summary Report for Spring 2014 the averages for each question range from 2.72-2.87. The difference between these numbers is so minute that there is no way to normalize the numbers.

---

3 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
4 Ibid.
5 Ibid.
6 Ibid.
7 Ibid.
8 Ibid.
10 Appendix A
OID’s reports also rely on standardized math definitions for their reports. The Department summary report lists the number of students who responded, the average, the median and standard deviation. OID does have some suggestions on how to interpret these standardized mathematical statistics. First, they suggest considering the sample size before taking the evaluations at face value, evaluations with less than two-thirds response should be viewed hesitantly and a number of 8 to 10 students responders is a minimum. 

Additionally, they suggest that these average ratings can be viewed on an absolute scale or in comparison to other courses/instructors. They use the example of a course with a mean average of 6.5 on a 9 point scale being above average, however in this scenario 70 percent of similar courses are also rated at above a 6.5 average. OID suggests that the course is therefore, “judged positively, but students were not particularly excited by it.” They also suggest looking at the standard deviation to gauge student’s variability of responses, still looking at a 9 point scale, a standard deviation of 1.0 is considered normal, while something like a 3.4 is indicative of a problem within the course. These deviations directly relate to the average course ratings given, so if the average is 7 then most of the students rated the course a 6, 7, or 8. They also suggest however to look out for grouping of students, which could indicate that two different large groups of students with different viewpoints. Lastly they suggest that course evaluations should not be used as the only measure of teaching ability because often they are influenced by outside variables. One of the biggest problems with using standard deviation, is with a 1 to 3 scale there is no real way to tell if there is a large deviation among students. On the Spring 2014 Department Summary Report the standard deviations range form .38-.57. These deviations are so diminutive and microscopic that nothing can be inferred or interpreted from the numbers. Also the median number is 3 on every question. These are standardized mathematical statistics but they are pretty much useless on such a small level.

The University of Oregon also looked at the statistical validity of student evaluations, citing report that suggested that a 58 percent response rate was necessary for validity in a class under 20 students, and a percentage of 35 percent for courses with 50 or more enrollees.

---

12 Ibid.
13 Ibid.
14 Ibid.
15 Ibid.
16 Ibid.
17 Ibid.
18 Ibid.
19 Ibid.


Storage

Student evaluations (both digital and paper) are stored in a plethora of places across campus. Most of the paper forms were given to professors and many still keep them on file. With the switch over to electronic student evaluations, the electronic copies of student evaluations are stored on servers around campus. The Student Evaluations are linked through each individual student’s MyUCLA account which leads them to the evaluation form, which directly deposits into OID’s database. From there EIP directly uploads the professor’s student evaluations to their MyUCLA accounts.

Previously these evaluations were given the Evaluation Coordinator on a CD and he or she would upload them to each professor’s account. Those CDs are still stored at Elizabeth Kalbers’ desk (IS’ Evaluation Coordinator); the CDs go all the way back to 2005. Now the Evaluation Coordinator receives EIP’s reports from box.com and then uploads the individual professor reports to the GSEIS server. Concurrently professors are expected to keep copies of their evaluations for future use. When being considered for promotion, they are asked to provide these evaluations to be compiled into a dossier; however if they do not have all copies requested they can request them from Marilyn Salinger, who retrieves them from GSEIS’ server. They also are retrieved from the server for Department Audits, which happen every eight years. The current system used to compile dossiers is by hand, however UCLA is working on a system called OPUS, which will digitally compile these dossiers and automatically pull evaluations reports from servers directly.

Access

Access to student evaluation information is restricted to personnel. The Department chair is not given direct access to any of the student evaluations but can request it. Professors only get access to their own evaluation information, and students do not have access to anything but the evaluation form they fill out themselves.

Both student and professor access is mediated through their MyUCLA accounts. In order to access, answer evaluation questions and deposit the completed evaluation into OID’s database, students click on the link provided in their

---

22 Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, March 1, 2016.
23 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
25 Ibid.
26 Ibid.
29 Ibid.
30 Ibid.
31 Elizabeth Kalbers, Interview by Rebecca Fordon and Dominique O’Neill, Jan. 27, 2016.
MyUCLA account or email; the link redirects and forwards the user to OID’s database. After submission the user is redirected back to their MyUCLA account. Individual instructors have access only to their own compiled individual professor reports, which they access through MyUCLA. The evaluation coordinator on the other hand has to accesses all instructor reports through the department administration server housed at the Education Technology Unit (ETU) on campus, which restricts access to the two coordinators for each department and Marilyn Salinger. Access is also being expanded to OPUS UCLA’s digital dossier system, which will automate compilation of previously manually compiled dossiers, and access evaluation servers to pull information directly. This lack of access, especially for students, will be addressed later in the analysis. It is worth noting that some Universities have provided students access to evaluations, University of Oregon provides students access to paper evaluations in a database, while the digital evaluations are available through “DuckWeb,” which requires a student login.

There have been some attempts at other universities in the UC system to expand access. The example of UC Berkeley is instructive here, as they have had the most open deliberations of the issue of access to student evaluations, although ultimately deciding against access.

In 2008, the Berkeley Division of the Academic Senate initiated a campus-wide Task Force on Teaching and Evaluation, to make recommendations about evaluation practices. As part of that process, in February 2011 a steering committee on online evaluations issued a report recommending that UC Berkeley “move toward a more open policy of sharing access to end of term course evaluation data.” The basic argument was that such data was already being shared via third-party sites, that UCLA could better ensure integrity of data that it administered, and that UCLA has an interest in “inform[ing] the campus community of the quality of instruction.” Specifically, the steering committee recommended that students receive access to both quantitative information (i.e., ratings) and new student-to-student questions (qualitative and quantitative questions that students could propose and that other students could answer).

________________________________________
32 Jessica Hoover, Interview by Rebecca Fordan and Dominique O’Neill, March 1, 2016.
38 Ibid., 4.
39 Ibid., 4-5.
The pushback from the faculty community was immediate and vigorous, with the Berkeley Division of the Academic Senate, the Berkeley (DIVCO), the Committee on Budget and Interdepartmental Relations (BIR) and the Committee on Educational Policy (CEP) all submitting written objections to the proposal for student access to evaluation data. These groups disagreed with the premise that allowing student “shopping” for courses was an appropriate goal for student evaluations, instead suggesting that student evaluations should be directed solely towards improving instruction. As to the legal issues of student and employee confidentiality, the groups had little response other than to say the law is unsettled and that UC Berkeley should be aggressive in protecting confidentiality.

**Circulation**

Outside the student evaluation production process, very little student evaluation information is circulated. The production circulation starts with students filling out the evaluations. From there quantitative and qualitative data is deposited into OID’s database and compiled and sent to both the evaluation coordinator and individual professors. Then, the coordinator deposits the individual evaluations into the administrative server where it is accessible by both the coordinators and Marilynn Salinger. Those then will get placed into a dossier when a professor is up for promotion and are circulated into a review board as evidence of a professor’s teaching ability.

**Visualization**

Summary information for each quarter is already contained in the rundown reports, which provide better visualization than tables of data. However it is not easily accessible or comparable to data over the course of a couple years because of its file format. Because of its dissemination in PDF this data can not be used for any other purpose than the one it was compiled in. This causes a problem because the information is valuable but not usable. There is also the problem of ambiguity in regards to the data that students produce through these evaluations. GSEIS’ current student evaluation form is form E, which was produced by the Education Department and presents the students with a 1 to 3 scale for each question. For example, one question is “Was the purpose of the class clear?”; the student is instructed from select 1 (“unsatisfactory”), 2 (unlabeled) or 3 (“excellent”). This scale is causes a lot of ambiguity, especially when the values are averaged, as all of the averages hover around the same number as previously illustrated in the standardization section. Both of these problems with visualization were corroborated when we interviewed professors about how they use student evaluations. This information is both extremely valuable and useful, yet it lacks the ability to be used to its full potential because of the form it is presented in.

---

40 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.


In terms of being able to use data in new formats there are a number of options. CSV is importable to spreadsheet software, though it sometimes requires the RFC guidelines that went into creating it.\textsuperscript{43} Microsoft Excel has a number of useful capabilities for consolidating multiple spreadsheets into one sheet.\textsuperscript{44} This would allow for the visualization of these data sets over multiple years, which would cut down on whiplash from moving from sheet to sheet. It will also allow for the ability to map out these averages over the course of a couple of years into charts for better visualization of trends, such as dips in performance.

\textit{Long-Term Preservation}

In our research of student evaluations we did not find any established long-term preservation methods. Most departments had a somewhat ad hoc policy to storing paper and digital copies of evaluations. Paper evaluations especially seemed to be left up to the purview of the professors. Digital storage was not viewed with any concern, which is a problem because most of the reports in the department are stored in PDF, which is subject to file format obsolescence. By additionally storing instructor report files in CSV, it would reduce the risk of file format obsolescence.

\textit{Work-Flow Analysis}

The work-flow of student evaluations through UCLA and the IS Department is summarized in the above section on context, and is reiterated in the below chart.

\begin{center}
\includegraphics[width=\textwidth]{chart.png}
\end{center}

Note that there are at least five transmission points of data. Each of these


transition points presents an opportunity for recontextualizing and for different viewpoints to be injected into the process. For example, when OID is compiling information from their database and creating summary reports, they are making decisions about which information (such as means and medians) will be useful to administrators and instructors and they are excluding other information from those summaries (such as a list of all scores for any particular question). The format in which they present these files also suggests to the user that the information being provided is all that is needed, and further data manipulation is not particularly necessary.

In the IS Department there is also an example of these different viewpoints. Despite being an Information Studies department, most people in the department were unaware of the other files collated by OID. Only certain files are transferred for backup, elevating their value above other student evaluation reports.

Similarly, when professors compile evaluations from their personal records to include in dossiers they compile for purposes of promotion, they may have an opportunity to provide context for those evaluations within the rest of their dossier.

**Governance**

**Institutional warrant**

Student evaluations are currently one of the primary (and mandatory) methods of evaluating faculty in the UC system for purposes of tenure and other advancement.

Student evaluations got their start at UCLA in the late 1960s and early 1970s. In 1969, then UC President Charles Hitch issued a memo to the UC Chancellors and members of the Academic Council, suggesting that opinion of students should be taken into account when assessing teaching effectiveness. 45 UCLA then formed a Task Force on the Evaluation of Teaching, which recommended that “an ongoing process of regular evaluations be conducted emphasizing the facilitation of faculty growth [and] the need for feedback for the development of UCLA’s instructional programs.” 46 Since then, UCLA’s Office for Instructional Development (OID) has conducted student evaluations through its Evaluation of Instruction Program (EIP). 47 For some period of time in the 1970s, these evaluations were even made available to students for their use. 48 Beginning in fall 2011, EIP has gradually transitioned to online student evaluations. 49

46 Ibid.
47 Ibid.
48 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
49 http://www.senate.ucla.edu/programreview/documents/SelfReview_OID_NoAppendices.pdf;
Student evaluations are also a major criterion in faculty advancement. Appointment and promotion decisions for UC academic appointees are generally made by an academic review committee that is nominated by the Committee on Academic Personnel or other committee. In the Information Studies department, the academic review committee consists of all department faculty. One of the four criteria on which the review committee must evaluate candidates is teaching (the other three criteria are professional competence and activity, creative work, and University and public service). The review committee must consider “clearly demonstrated evidence” of teaching, as stated in the Academic Personnel Manual (APM):

Clearly demonstrated evidence of excellent teaching is an essential criterion for appointment, advancement, or promotion. Under no circumstances will security of employment be conferred unless there is clear documentation of outstanding teaching.

Further, the review committee must “clearly indicate the sources of evidence on which its appraisal of teaching competence has been based.” Among the “significant types of [such] evidence” are students’ opinions, and “[a]ll cases for advancement and promotion” should include “evaluations and comments solicited from students for most, if not all, courses taught since the candidate’s last review.”

This policy is consistent with the following June 5, 1972 resolution by the UCLA Legislative Assembly, which is cited in an appendix to the APM on documenting effective teaching:

It is essential to the evaluation of instructional quality and impact that candid, non-selected and reasonably complete student opinion on teaching effectiveness be obtained for all courses and instructors. Student opinion, in writing, should be regularly solicited for all course offerings, and each department or school should devise its own procedures to this end. Reasonable uniformity and consistency in procedures within each department should be maintained, but it is recognized that differences in subject matter and methodology between departments make it unreasonable to specify a University-wide format.

Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.

53 Ibid., 210-3-c-1 (emphasis added).
54 Ibid.
55 Ibid.
Stakeholders

We identified a number of stakeholders for the information produced from student evaluations, each with different interests in the information. In instances where stakeholders currently have no access to student evaluation information, we identified how the information affects them and we also identified their interests in expanded access.

University (UCLA). Regulations from both UC and UCLA mandate that review committees consider written course evaluations as part of the evaluation of academic appointees. The stated rationale is that such evaluations are “essential to the evaluation of instructional quality and impact.” Notably, UCLA (under the auspices of the Office of Instructional Development (OID)) is the only stakeholder that has access to the full evaluation database.

Department (IS). The Department is also interested in evaluating instructors, but from a standpoint of identifying and responding to issues, as opposed to for purposes of advancement. In addition to the academic review committee’s formal use in the review process, the department chair looks at the quantitative evaluations to determine if there are any issues for any particular professor and, if so, then looks at the qualitative comments. If there were any issues, he would speak to the relevant professor.

However, the evaluation information the IS department chair receives is limited in both form and content. As described above, every quarter the department chair receives (through the evaluation coordinator, Elizabeth Kalbers, reports on all instructors for that quarter, but receives no longitudinal information comparing either instructors or the department over time. OID does make other summary information available to department heads (again through the evaluation coordinator), but due to some quirks in the process, the current IS department head had not seen any of these summary reports before we made him aware of them. Further, even these reports provide information only for the quarter in which they are produced. Thus, the department’s stake in student evaluation information is in some ways restricted by the form in which OID provides the information.

Faculty. Instructors have a vested interest in both the power of student evaluation information to affect their employment as well as the use of student evaluation information as a tool to improve teaching quality and the curriculum. In speaking to faculty, we discovered the interest in student evaluations for teaching purposes sat somewhere behind their importance in the review

59 Ibid.
60 Ibid.
They found qualitative information (comments) to be more informative than the quantitative information (numerical ratings), but overall the comments did not provide a great deal of new information or contained mean spirited comments containing no constructive criticism. Professors indicated and insisted they usually already knew about any problems (both from students’ in-class reactions and from their own instincts).

Instructors receive a report every quarter that contains quantitative and qualitative information for each of the courses they taught that quarter. However, the reports contain no longitudinal information.

Similar to the department head, faculty have expressed interest in longitudinal information that could help them to identify trends in quantitative evaluations over the course of several years, which they could then connect with any changes they had made in the course or changes they had made in their teaching style.

**Current Students.** Although students do not currently have access to formal student evaluation information, they nevertheless have a strong interest in producing that information by completing student evaluations, and therefore making their opinions heard. Completing student evaluations provides students with the apparent power to influence advancement of faculty, revisions in teaching methods, and revisions to curriculum.

Students also have an interest in student evaluation information as a way to determine which courses to take. This interest is not being fulfilled by OID’s formal student evaluation information since students do not have access to it. To fulfill this need, students then must turn to other forms of information such as speaking to their classmates in-person and in online communities like Facebook. Other avenues that seek to fill the demand for information include third-party evaluation systems such as bruinwalk.com and RateMyProfessors.com.

**Prospective Students.** Prospective students also have an interest in the quality of teaching instruction and the quality of the curriculum as criteria for differentiating between programs. Like current students, prospective students have no access to formal student evaluation information. They also have limited access to conversations with current and former students. They are therefore limited to third-party evaluation systems such as bruinwalk.com and RateMyProfessor.com, as well as less granular ratings such as those from U.S. News and World Reports.

---

64 Student Discussion, Data Informatics, February 9, 2016.
65 Ibid.
The public. Because the public funds the UC system, the public also has an interest in the quality of the employees in the system, particularly in relation to their salaries. This interest is currently addressed through UC and UCLA policies and procedures for instructor review and advancement. The public has no direct access to student evaluation information.

Legal and Regulatory Restrictions on Access

Disclosure of student evaluations is governed by a mixture of state and federal law. The California Public Records Act of 2004 (CPRA) addresses the disclosure of employee information (such as information tied to a particular instructor), while the Family Educational Rights and Privacy Act (FERPA) addresses disclosure of student information.

CPRA restricts the disclosure of employee information by public institutions. Although CPRA makes publicly available all records relating to the conduct of the public’s business (including the conduct of a public university such as UCLA), the statute exempts from disclosure a number of categories of information, including “[p]ersonnel, medical, or similar files, the disclosure of which would constitute an unwarranted invasion of personal privacy.”70 There is little case law on the topic in California. However, some California universities have concluded that disclosure of student evaluation material would not constitute an “unwarranted invasion.” Arguments in favor of disclosure include: “(1) faculty should have no(or limited) privacy rights concerning information about the quality of service they provide at a public university; (2) there is a strong public interest in students being informed about the quality of faculty teaching; (3) teaching evaluations are merely collections of student opinion, which students are free to share publically [sic]; and (4) similar information is widely available through websites such as ‘Rate My Professor’.”71

The disclosure of personally identifiable student information is governed by FERPA, which provides that institutions may not disclose such information without written consent.72 However, there are a number of exceptions that nevertheless permit disclosure without consent, including disclosure “to organizations conducting studies for, or on behalf of, educational agencies or institutions to: (A) Develop, validate, or administer predictive tests; (B) Administer student aid programs; or (C) Improve instruction.”73 Again, the law

---

66 Although one of the authors of this report is a former lawyer, nothing in this report is intended to create an attorney-client relationship and this report does not constitute legal advice. The report can be viewed as a summary of legal principles and resources, but the authors direct readers to consult with legal counsel with regards to any legal questions.


69 California Gov. Code § 6253.

70 California Gov. Code § 6254(c).


73 34 C.F.R. § 99.31(a)(6)(i).
is not clear on how this applies to student evaluations and whether information that has been stripped of identifiers (such as the de-identified data produced from online course evaluations) can be considered “personally identifiable student information.”

**Impact on Record-Keeping Norms and Policies**

Although UCLA appears to have no formal policies on disclosure of student evaluation information, no one we interviewed could provide us with any information from the OID student evaluation database beyond departmental averages that contained no information specific to a particular instructor or course.

The record-keeping practices are also influenced by the timeline for faculty reviews, which tend to occur every five years. We also discovered no records destruction policy at the University or department level. Other public universities in California also restrict public and student access to student evaluation information.

**Transparency**

The process of producing, storing, and transmitting student evaluation information is fairly transparent. OID and EIP provide information about their process on their website.

However, less transparent is the history of decisions that resulted in this process. We obtained much of the information in this report through interviews and through extensive research, but the information was not readily available from one place like information on the professor evaluation process. Further, there is information that is still opaque, such as whether the Academic Senate or other decision-making bodies at UC and UCLA have ever considered providing open access to student evaluations and what their rationale are for keeping that information private. Furthermore the actual reports themselves were unattainable due to privacy concerns, making even the layout of those reports a mystery to most students.

Some discussion on these topics arose in the early 2000s, when bruinwalk.com was becoming popular, but the discussion appears to have petered out before any consensus was reached. And in contrast to other universities in the UC system such as UC Berkeley, the UCLA Division of the Academic Senate does not appear to have deliberated upon the issue of openness of student evaluation data.

---

Prospective Analysis

Trends

In formulating our recommendations, we considered several trends that are likely to influence the future use of student evaluation data. Each of these is discussed below, along with its implications for future use.

Social media. Due to Information Studies department size, gathering data on students and prospective courses is often easier than at an undergraduate level. It also allows for the formation of things like Facebook groups for MLIS students. However, there are significant obstacles to information gathering for students. Firstly, often classes are only offered once a year, making comparisons difficult between students. Also, the department is in a period of transition, making the number of new classes offered more numerous than in previous years.

With the formation of things like Facebook groups and other social media platforms for discussion, there also feature some flaws. Firstly, it assumes that every student has a Facebook account or even is invited to the group. As was shown in IS 289, many students in the class were unaware that such a group even existed. Second, the group as it exists today only serves to connect MLIS students of the class of 2017. This leaves out an entire group currently inhabiting our department with valuable information in which to share.

Open data and aggregation of data. Open data is a movement whose goal is to increase accountability and find meaning out of data. However, open data can also be misused to draw incomplete conclusions or remove privacy. Open data is a response to the trend of digitization of information and trying to extrapolate data from this information.

The trend of open data correlates with an increase in uses of open data standards like CSV. This standard is an open format, which is both machine and human readable, making it a good choice for open data evaluations. However, CSV, like open data, has its limits, sometimes needing extra tools to make the data useful.

With the open data movement, an increase in the aggregation of large amounts of data has been increasingly apparent, both in private and open forms. The aggregation of large amounts of data in databases is the basis of the big data movement. This is best shown through LA city’s aggregated open data dashboard, with open city data. The open data implemented in a dashboard

---

78 Student Discussion, Data Informatics, March 8, 2016.
79 Student Discussion, Data Informatics, January 5, 2016.
81 Ibid. 3-4.
82 Ibid. 3-4.
83 https://performance.lacity.org
not only illustrates this trend, but also points to the manipulation of this data for both political and environmental use. It also points to the increasing belief that this data can effect change.

Both of these trends may have implications for student evaluation information. We are already seeing student evaluation becoming aggregated into dashboard applications like OPUS, who aim to help gather data for promotions dossiers of UCLA faculty. We may also see further pushes for openness of data, particularly as both state and federal law begin to mandate greater accessibility, along with it being a good political tool.

**Demand for consumer reviews.** The past few decades have seen an increase in the prevalence of online consumer reviews, such as on amazon.com and yelp.com. This trend is already beginning to carry over into student evaluations, through websites like RateMyProfessor and bruinwalk. As students, we have an unique perspective on the usefulness of student evaluation sites geared towards garnering our business. The problem with sites like RateMyProfessor, especially for the Information Studies department, is a lack of data. Some Professors, like Professors Leazer and Drucker have no reviews whatsoever, while some like Professors Noble and Furner have only 2-3 reviews. Few professors in the IS department have more, but no professor has more than eleven reviews. Bruinwalk on the other hand has the same problem: there is very little information on each professor and because they split reviews based on course and then professor there are even fewer reviews per page. Even if there were a plethora of information on either of these sites, like there is for the undergraduate classes, its utility and usefulness may be limited. For example, in my experience as an undergraduate, most of my peers used RateMyProfessor or similar sites to choose a class because sparingly because we found the reviews were often not particularly helpful, instead we tended to use the same techniques of social interaction to learn about professors and course, despite the school’s large size. Despite the fact that consumer reviews have become a widespread trend, most people will trust someone they know has used a product, like a friend, as opposed to an anonymous source online. There is still tremendously more trust if the potential user knows for sure that the reviewer has actually used the product and like it. These problems are inherent to almost all consumer reviews, including Student Evaluations, which will be further explored in the next section on trust, validity and bias.

**Points of resistance**

**Trust, validity, and bias.** One of the significant limiters for any expanded use of student evaluation information is the lack of trust in the quality of the data. We found this concern present in both faculty and students, although to different degrees for quantitative and qualitative evaluations.

For example, some faculty members questioned whether student ratings are really effective evidence of teaching ability, and several said that they pay little attention to numerical scores. This concern is born out in the academic literature on the issue. In their article “An Evaluation of Course Evaluations,” Philip B. Stark (a professor of statistics at UC Berkeley) and Richard Freishtat (of UC
Berkeley’s Center for Teaching and Learning) argue that student evaluations do not accurately measure teaching effectiveness. They summarize the available literature as concluding that student evaluations instead correlate to a wide variety of other factors including: students’ grade expectations, their enjoyment of the course, and characteristics of the instructor such as physical attractiveness, gender, ethnicity, and age. Further, they note that broad ratings questions about “curriculum design, effectiveness, etc.,” appear to be most influenced by factors not related to learning.

The validity of evaluations is further undercut by low response rates, particularly after the switch to online evaluations. Press reports from 2013 estimated that response rates for online evaluations were approximately 42 percent, as compared to paper response rates of 75-85%. Anecdotal evidence from faculty in the IS department suggests online evaluation response rates can be as low as 20-30% for some classes (as compared to response rates of 80-90% for paper evaluations). However, at least some studies have concluded that, although response rates differed between paper and online evaluations, the scoring patterns are highly consistent between the two. OID reports that its own studies have confirmed this finding.

University of Oregon cites the findings of eighteen studies on student evaluations accuracy from paper to digital, claiming that of the eighteen, fourteen reported no difference in accuracy and 2 reported slightly higher averages from online evaluations. They also claim that by moving to online evaluations students contribute more positive and useful comments than paper evaluations. However, they do accept that return rates are lower online, barring messages and rewards, the difference totalling 9%. However they claim that studies have found incentives can boost response rate from 7-25%. University of Oregon who does use incentives, claim their response rate is 78-79% (65% who fill out the form and don’t decline), however they have paper evaluation percentage numbers to compare to.

---

85 Ibid.
89 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
91 Ibid.
92 Ibid.
93 Ibid.
Finally, the very nature of the ratings also makes them difficult to interpret as averages. Stark & Freishtat explain that student evaluation ratings are “ordinal categorical variables,” which means the ratings fall in categories in a natural order (worst to best), but are labels rather than values and thus are not susceptible to averaging. Stark & Freishtat illustrate this point with a joke: “Three statisticians go hunting. They spot a deer. The first statistician shoots; the shot passes a yard to the left of the deer. The second shoots; the shot passes a yard to the right of the deer. The third one yells, “we got it!”

This all corresponds to the hesitance of faculty to rely on quantitative data (ratings), and their greater interest in qualitative comments. The problems discussed above further suggest that faculty tolerance for the current structure of online evaluations may not last, particularly if student evaluations begin to become publicly available or widely shared (such as through OPUS).

There is also no regular review of questions to determine if they are accurately measuring what they are intended to measure. The current questions have not changed for at least the past ten years. Instead, quality is controlled by ensuring that students are enrolled in a course before they are permitted to comment (through a portal to the registrar’s information).

To generate more trust and limit bias, the Department may consider altering the questions on its student evaluations. For example, fewer broad questions may provide less opportunity for non-teaching-related factors to influence students. Further, different visualization methods (and elimination of reliance on averages) may provide a better picture of the evaluations. For example, a scatter plot would show whether ratings tend to cluster around a high rating or a lower rating.

As we have discussed in the trends section, consumer reports have a problem of lack of trust that the reviewer actually owns or uses the product. They are susceptible to tampering to improve ratings or mischief. One of the biggest values in student evaluations given by the school is the fact that they ensure that the reviewer of the class has actually been registered for the class. In order to fill out an evaluation the student has to login into their MyUCLA account and be registered for the class. This ensures that the trust and validity of the reviews is extremely high. Unlike RateMyProfessor or bruinwalk, the user can be assured to the fact that the student was a member of the class. For example Professor

95 Ibid., 6.
97 Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, March 1, 2016.
98 Student Discussion, Data Informatics, February 9, 2016.
Blanchette’s bruinwalk page features a review written by him; there are no safeguards in place to restrict access.99

The Department could also consider increasing incentives for producing quality evaluations. Currently OID provides for incentives by tying a participation grade in MyUCLA gradebook to the filling out student evaluation forms.100 However this requires a Professor to use the MyUCLA gradebook. The Education Department has suggested that since most students have devices in class already, that Professors set aside time for students to fill out the forms in class. This is much like the system that used to be used with paper forms and the hope is it will increase participation.101 The University of Oregon has an incentive system that places a grade hold on student’s accounts, which means that if students do not fill out or opt out of student evaluations, their grades, unofficial and official transcripts are unavailable until after the semester grading submission deadline for professors has passed.102 Considering that instructors seem to prefer qualitative comments, the department could consider whether simply to remind students to fill out both sections or allow them to leave only qualitative comments if they prefer.

**Privacy and anonymity.** In looking at the privacy of student evaluations two groups’ anonymity must be addressed: students and professors. We have already looked at the legal implications of student evaluations and access in the Legal and Regulatory Restrictions sections, here we will look at the social implications of privacy. According to OID and EIP there is no way to tie students to the reviews they write, the information itself is never entered into the database.103 If this is true then these reviews are fully anonymous, which like most online reviews makes it hard to tie back to a specific person. However in a student survey of IS 289, about 76 percent stated they wanted access, but some worried about student anonymity being maintained, suggesting that some of the comments could give away information on who wrote the review.104 Speaking as a student, I would think it would be more of a concern of students that they might give information in comments that could give away their identity to professors. If there is still a problem of anonymity, simply informing the students that their evaluations can be accessed would be a good step, and making these reviews only available to students through a valid login.

There is of course going to also be a problem of data quality tied with anonymity; if the reviewer can say whatever they want without consequence, there can be a loss of good data. The professors we interviewed often said students can be very harsh in regards to what they say in their reviews and they often do not take

---

99 Student Discussion, Data Informatics, February 9, 2016.
100 Marc Levis-Fitzgerald, Casey Shapiro, and Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, Feb. 1, 2016.
103 Jessica Hoover, Interview by Rebecca Fordon and Dominique O’Neill, March 1, 2016.
104 Data Informatics Class, Survey, March 1, 2016.
grating or cruel reviews seriously.\textsuperscript{105} Instead of receiving constructive criticism, reviewees can often receive personal attacks which leads to data quality loss. Honestly the only way this can be handled, is the way it is currently being handled, requiring the student to already be registered for the course. If the student evaluations were not anonymous, the likelihood that any student would willingly fill out evaluations is unlikely, especially with how low response rates already are. Perhaps the only way to increase data quality would be to increase response rates. As evaluations are now, most professors see the only people who respond are people who loved the class or hated the class, there is very little in between.\textsuperscript{106} Perhaps by implementing time in class to fill out evaluations, much like was previously done with paper evaluations, the response rate and quality of data could increase.

Finally, professor anonymity should be protected. When the authors were collecting data, the only report we were able to gain access to was the department summary. Anything with individual professor information included was off limits to us. Professors should of course be consulted before their information is opened to students, however one would think that they would prefer the evaluations of their teaching be accurate as opposed to open to anyone.

**Conclusion**

Student evaluations are of course a sticky subject for both those who evaluate and those evaluated. They also are a huge mine of information, which is why companies like RateMyProfessor exist. This report hopes to help clarify and demystify the process behind student evaluations and help point out potential pitfalls and problems inherent in student evaluations. The next section will present our recommendations in hopes of improving how the data can be used for the Information Studies department.

**Recommendations**

**Immediate**

- **Request department-wide .csv files from OID.** Hopefully one of the easiest recommendations to implement will have to do with changing the file formats in which EIP reports are being disseminated. One of biggest problems with the pdf files that EIP current hands out to departments is they can not be used to look at any data besides the specific data provided by EIP. Looking at trends over the course of multiple quarters or years is not currently feasible with the pdf format. Hopefully with the transition to CSV files, the department will make use of the reports that EIP compiles, instead of ignoring them completely. Perhaps asking for .csv


that contains information for the past seven years will also help make the gradual cumulation of data more manageable.

- **Provide access/save other evaluation data.** Assuming that the short-term solution of evaluation files now in .csv format, provide access/save the other evaluations OID provides, currently they are left on a box.com account, but not used or saved anywhere else.

- **Provide visualization tools to faculty and department head.** Again assuming that the short-term solution of evaluation files now in .csv format it should be trivial to insert data into some tool, such as Microsoft Excel, that can provide different visualizations, for example to visualize a scatter plot over the course of a number of years.

**Mid-Term**

- **Consider providing mechanism longitudinal student interaction.** Consider a student-to-student question section on course website, where questions remain for future years.

- **Consider revising rating scale.** The scale that currently exists in GSEIS’ student evaluation form is on a 1 to 3 scale. This provides very little granularity and obscures the data collected. We suggest collaborating with Education department to create a new form with either a 1 to 5 scale or 1 to 9 scale. This will increase readability of the data and hopefully reveal more transparent areas of improvement. In revisiting the scale used on the student evaluation, perhaps the form and questions asks should also be revisited for calibrating student responses better.

- **Work with student organizations, such as the Student Government Board for Information Studies, to develop course selection tools.** Given that the impetus for access to student evaluations is the need to select courses, providing reliable tools could alleviate the need for such access. These tools could include other data (such as course syllabi information) and could also incorporate a discussion board that is moderated by students so as to attempt to eliminate the effect of bias.

**Long-Term**

- **Advocate for improvements to evaluation process.** If the justification for restricting student access to student evaluation information is that the information is not reliable enough, then that should also be cause for improving the evaluation system so that the information is more reliable such that it can be shared with students. Revisiting the form that students currently use and improving the questions could help ensure better data quality.

- **Consider ways of increasing response without receiving bad data.** This report has outlined a number of incentives suggested or used by other programs/schools. Looking at ways of increasing user response can
possibly increase the amount of reliable and balanced data going into the process.