Final Report of the Task Force to Create a New Student Ratings Form for California State University Channel Islands

Task Force to Create a New Student Ratings Form (SRF)
"Adams, Virgil" <virgil.adams@csuci.edu>
"Bleicher, Bob" Bob.Bleicher@csuci.edu
"Buchanan, Merilyn" Merilyn.Buchanan@csuci.edu
"Ed Nuhfer" ed.nuhfer@csuci.edu (Chair)
"Elliott, Jesse" jesse.elliott@csuci.edu
"Furmanski, Matthew" matthew.furmanski@csuci.edu
"Renny.Christopher" Renny.Christopher@csuci.edu
"Smith, Peter" peter.smith@csuci.edu
"Wood, Gregory" gregory.wood@csuci.edu

Table of Contents

Need for a New Student Ratings Form....2
Student Ratings and University Policy....2
Background information....3
Opportunity to Pilot an Online Ratings System....4
Uniqueness of California State University Channel Islands....5
Current and Applicable Literature....6
Global Items....7
Examination of the Student Ratings Form, its dimensions and its items....9
Calculated Global Item....15
Open-ended Items....15
References Cited....16
APPENDIX A Minimal Requirements for an Online Ratings Delivery System....19
APPENDIX B
PART 1. Screening Criteria....21
PART 2. Some Guidelines for Creating Good Items....22
APPENDIX C Proposed Ratings Form for CSUCI....24
APPENDIX D Reasons for Affective Wording on Student Ratings Forms....25
APPENDIX E Some options considered for summary reporting 29
APPENDIX F Preliminary Results from Spring 2008....31

Report of the Task Force to Create a New Student Ratings Form for California State University Channel Islands

Need for a New Student Ratings Form

The CSU Channel Islands campus adopted the Instructional Assessment Form X from the University of Washington, (see this and other forms at http://www.washington.edu/oea/services/course_eval/forms/index.html) as its first student ratings form, primarily to meet the policy requirement to procure a ratings instrument with some expediency. The form is a scannable paper form, copyright date 1995, which is mailed to and scan-processed by the University of Washington, then returned to this campus. The expense of the paper forms and processing are at least $25,000 a year and increasing rapidly.

Because of a decision by the CSUCI Academic Senate at the time the form was adopted, every "Form X" used at CSUCI must be run manually through a laser printer to black out items the Senate chose to censor before the forms can be distributed. Maintaining the current system is untenable.

In fall of 2007, Terry Ballman, Chair of the Academic Senate requested that the Faculty Development Advisory Committee (FDAC) at Channel Islands find a replacement for the current ratings form. As work commenced, confusion arose as to what committee should have oversight—FDAC or the Faculty Affairs Committee (FAC). The solution was creation of a Student Ratings Form Task Force in December of 2007, with membership from the Faculty Affairs Committee, the Faculty Development Advisory Committee, the Associate Vice President of Faculty Affairs, and the Director of Faculty Development, who became chair by the second meeting by consensus of the task force.

Student Ratings and University Policy

Channel Islands is bound by the Collective Bargaining Agreement (CBA) of May 15, 2007, between the California Faculty Association (CFA) and the California State University (CSU). The current policies appear in the recent Joint Committee Report (2008) on Student Evaluations of Teaching.

According to Article 15.15 of the agreement, “Written student questionnaire evaluations shall be required for all faculty unit employees who teach. A minimum of two (2) classes annually for each faculty unit employee shall have such written student evaluations. Student evaluations shall be conducted in classes representative of the faculty unit employee's teaching assignment. The results of these evaluations shall be placed in the faculty unit employee's Personnel Action File.”

The CBA stipulates that these evaluations must be anonymous, and either quantitative (meaning survey data that can be expressed numerically) or a combination of quantitative and qualitative (normally implemented either through the use of open-ended questions, or through the provision of opportunity for students to write comments).

Background information

The volume of literature written about student ratings is immense—larger than any other single topic in higher education. Cashin (1988) noted that over 1300 articles on the topic existed in 1988 and the number today is more than double that figure.

There are two very different kinds of student ratings: *formative* and *summative.* Formative evaluations ask detailed questions that provide a profile of pedagogies and strategies being employed. These diagnostic surveys yield information about teaching traits and practices that allow professors to make changes and improve instruction. Formative evaluations are given during the course, usually about mid-term.

Summative ratings are given at the end of a course and are used primarily to evaluate professors for personnel decisions that involve retention, promotion, tenure, and/or salary. The numerical scores that express the student ratings of professors on summative evaluations are sometimes called "student satisfaction" or "student opinion" ratings or "student evaluations." The use of the last term has fallen recently into disfavor. Scores arise from complex factors that include cognitive learning, teaching traits, superfluous attributes, and affective personal reactions that are products of both what happens in a class and what an individual student brings with him or her to the class in form of bias and motivation. The CSUCI Student Ratings Task Force was charged with constructing a new summative rating form for the University.

Our task force focused on particular research in which data from hundreds of summative and formative surveys were combined in meta-analytic studies. One purpose of these studies was to discover what teaching traits and instructional practices contributed the most heavily to *high summative ratings*. From distillation of literature and empirical data analyses, researchers discovered related traits and practices that grouped together into "teaching dimensions" (Feldman, 1998). Although different workers deduced and reported different numbers of dimensions, about half a dozen consistently appeared as the most important on all lists.

A similar meta-analytical approach was used in research that deduced which teaching practices were most important to *student achievement*. Student achievement was usually defined by concrete measures such as test scores and grades. The dimensions deduced as important to student satisfaction are similar, but not identical, to practices that produce student achievement. Several dimensions deduced as important to achievement have been validated experimentally (Hines, Cruickshank, & Kennedy, 1985; Schonwetter, Menec, & Perry, 1995). Recently, Pascarella, Seifert, and Whitt, (2008) studied a population of students to verify that these same dimensions are important to increasing retention and persistence in first-year college students.

We focused on this block of research because there are an infinite number of questions one can ask on a student ratings form. Knowing the most important dimensions allows creators of a student ratings form to focus on creating only items that relate to instructional traits of proven importance. In short, we wanted to try to collect only the data worth collecting for summative purposes.

In addition to two face-to-face meetings, most task force work took place online or between teams of members. The task force recognized their first priority as producing a core ratings form. Some members of the task force worked on an attempt to create a replacement rating form in 2006-2007, only to have their proposal fail to make Academic Senate approval, so they accepted their appointments to this 2007-2008 Task Force with some trepidation.

Opportunity to Pilot an Online Ratings System

Given economics and the proclaimed sustainability emphasis of the Channel Islands Campus, the task force decided to try to open opportunities to dispense and process student ratings through an online system. The Director of Faculty Development did a national query to discover where online systems are used successfully and how. In December, task force members and members from instructional technology attended an online workshop on online student ratings forms conducted by Dr. Peter Gold of CourseEval®. The demonstration was impressive and reception here favorable. Adoption of an online system would represent a major transition for any California State University campus.

CSUCI's specialists in Instructional Technology administration, largely Judy Swanson and Peter Mosinskis, are working to arrange a pilot test of CourseEval® at this campus. At their request, the Director of Faculty Development drew up a list of desirable criteria for an online system. These criteria may be found in Appendix A of this report. The efforts of the rest of the task force then had to be re-directed back to focus on the task of constructing the ratings form.

The Task Force formally considered four options, and the options and decision about each option follow.

Retain the present form	Rejected as both undesirable and untenable
Procure a commercial form	Rejected on the basis of being costly and the likelihood of getting a bad fit, but this option remains a possibility if all else fails
Procure a form from another institution	Rejected because of the likelihood of getting another bad fit and continued high costs
Create our own form	We chose to proceed to create our own form, in order to have one that fits our needs at this unique institution.

We proceeded to compose our student ratings form based upon the uniqueness of this campus and the most current and applicable literature.

Uniqueness of California State University Channel Islands

The key to matching a form with the needs of an institution begins by considering the institution's mission statement. The CSUCI mission statement is "Placing students at the center of the educational experience, California State University Channel Islands provides undergraduate and graduate education that facilitates learning within and across disciplines through integrative approaches, emphasizes experiential and service learning, and graduates students with multicultural and international perspectives." This provided keys that guided construction of our ratings form.

Ratings forms tend to be heavily teacher-centered, course-centered, learner-centered, (see Barr and Tagg, 1995) or a random mix that tends to give the form no focus. Our mission statement indicates that a *learner-centered* student ratings form probably best matches the needs of our institution.

The prevalent commercial ratings forms and forms created for use at other institutions (with perhaps one exception addressed below) are invariably designed for lecture-based classes. Our campus encourages a variety of pedagogical approaches in addition to lectures that involve interactive engagement. As a result, we needed a form that was *pedagogically-independent*, in which the ratings items could apply equally to lectures, cooperative learning, studio classes, experiential and service learning, case method, Freirean approaches, *etc.*

Different disciplines try to achieve competency in ways that translate into conceptual differences between what takes place in courses. Some are knowledge-rich, others process-rich, others skill-rich, and others awareness-rich. Because of our interdisciplinary emphases, a general ratings form based on items that are *discipline-independent* probably best serves our needs.

The Channel Islands campus has an additional signature character worth considering in design of a ratings form: its outstanding climate of creativity that nurtures not just positive collegial relationships among its constituents, but unusually positive ones. Much of the literature on student ratings reveals that poorly done evaluation or misuse of evaluation is destructive to a campus community. It can pit faculty against one another, faculty against administration and students against faculty. We felt that if this campus of risk takers can create a unique and positive campus, that we can also create an evaluative system in accord with our unique campus signature. In short, evaluation should be an event that offers learning in general, and not simply be an exercise in rating teachers with numbers. Further, any criteria used to rate faculty should be criteria that they can control, improve upon, and be supported to improve upon through faculty development. McKeachie (2007) hits a chord which resonates with the philosophy of this campus, and with which we created this student ratings form

> Systems of student evaluation of teaching should encourage students to think about their own educational experiences—to develop clearer conceptions of the kinds of teaching and educational experiences that contribute most to their learning. The student opinion form could, and should, be educational in the highest sense—helping students gain a better

understanding of the goals of education, stimulating them to think metacognitively about their own learning, motivating them to continue learning, and encouraging them to accept responsibility for their learning.

Current and Applicable Literature

We had the benefit of prior thorough reviews of literature from several sources that included the task force report of 2006-2007, the compilation by Nuhfer (2007), and additional newer resources that were not available to the task force of 2006-2007. The most critical literature to construction of our ratings form was that establishing the current legitimacy of the most important teaching dimensions. We found through Theall and Feldman (2007) and Abrami, Rosenfield, and Dedic (2007) and Abrami, d'Apollonia and Rosenfield (2007) that Feldman's original dimensions (see Feldman, 1998) remained valid. These dimensions listed in order of importance (from #1 as most important) to learning (from Feldman, 1998) follow.

Instructional Dimension	Ranked Importance with Learning Achievement	Ranked Importance with Satisfaction	Achievement % Variance Explained
Preparation and organization	1	6	30-35%
Clarity and understandableness	2	2	24 -30%
Perceived outcome or impact	3	3	15-20%
Stimulation of interest in content	4	1	10-15%
Encouragement and openness	5.5	11	<10%
Availability and helpfulness	5.5	16	<10%

Considering the above, our Task Force had the choice of constructing the form based on dimensions important to learning or dimensions important to overall satisfaction. The Task Force deemed learning as the more important of the two; thus our draft student ratings form is both learner-centered and learning-centered. We constructed ratings items that addressed the instructional dimensions listed above.

Two resources aided our drafting of items: Berk (2006) and Nuhfer and Dewar (2008-Appendix B part 2 of this report). The criteria for screening items (Appendix B part 2 of this report) are rigid, but these criteria had been not easily accessible in a single source until recently. These compilations were not available to the task force in 2006-2007, and indeed items on their draft form could not pass these critical criteria. However, some items on most commercial forms, including those that tout their psychometric reliability and validity, and the U of Washington "Form X," currently in use by our campus, can't pass these criteria either.

The Faculty Development Office purchased a copy of Berk (2006) for each of the Task Force members and for the Channel Islands Broome Library. Every item in the proposed form had to pass twenty-five criteria (Appendix B Part 1), most of which are listed in Chapter 3 of Berk (2006). All task force members were involved in drafting items and/or

screening each item to see that it could pass all the criteria. As one member of the task force noted regarding drafting items that could meet all criteria, "This is harder than I thought it would be...." Indeed, an individual will not create a good form alone. A committee that functions as one large brain, with all members supporting one another in proposing alternatives when items fail, is perhaps the only way to succeed.

The only other institution we are aware of to design a pedagogy-independent student ratings form is University of Minnesota (Langley et al., 2007). Their task force produced their ratings instrument by coupling six core items based on the important teaching dimensions, three open-ended prompts for students' written comments, and a number of additional items based directly upon the student learning outcomes specific to the course and the institution. Our task force considered use of learning outcomes as part of our summative form, but members were split on the wisdom of trying to do so at this time. This remains a possible option for the future.

The employment of the University of Minnesota's six items without student learning outcome items results in a form too short to yield good reliability, and one or more of these items might not pass the twenty-five criteria mentioned. We opted to create our summative rating survey based on several items for each of the teaching dimensions most associated with producing learning, supplemented by three open-ended comment items.

Global Items

Summative items called “global” do not obtain information under the category of any specific teaching dimension. Instead, they solicit a general overview of the satisfaction with the course experience. Typical global summative items include

"Overall, how do you rate this instructor’s teaching ability?”
“Overall, how do you rate this course?”
"How do you rate this course as a learning experience?”

The University of Minnesota opted to omit any global item from their new form, based upon bad experiences with the global item in previous years. They noted: "While the score of one global item can result in an efficient evaluation by personnel committees, the item tends to be relied upon as a single proxy measure of instructional performance.... Finally, there is little to no diagnostic value for a teacher who scores medium to low on this item compared to items that measure specific aspects of teaching (Arreola, 2006)."

The experience noted by the U of Minnesota is typified by misuse of global items at many institutions. Robert Leamnson, the late author of the popular book, *Thinking About Teaching and Learning* (1999, Stylus) disclosed

> Over 25 years I was evaluated by more than 100 groups of students. For about 18 of those years, I participated in annual reviews of my departmental colleagues and in those of all Arts and Science departments at the college level. The methods were virtually invariable. Whenever a rating form had a global question, it alone was considered and all else simply ignored! If, in the real world, you put a global question on the rating form...the global question is the end-all and be-all—the alpha and the omega.
>
> Robert Leamnson, November, 2005, personal e-mail communication to Edward Nuhfer

In contrast, Abrami, Rosenfield, and Dudic (2007) argue on the basis of factor analyses that "In brief, we recommend that only global items should be used for summative decisions." To be sure, this is a minority opinion. If it were not, universities would largely be doing summative evaluations with forms that ranged from one to three global items in length, and such ratings are unheard of. Later, these authors seem to offer a contradictory conclusion when they state: "An emphasis on student-centered learning has made traditional forms of student ratings of questionable relevance as a universal approach to judging teaching effectiveness" (Abrami, Rosenfield, and Dudic, 2007). In truth, long before interactive engagement pedagogies ever became common in college classrooms, the majority of researchers were consistent in recommending strongly against the use of student ratings as a sole basis for judging teaching effectiveness.

There are good reasons to regard direct global rating items, in particular, with suspicion for important personnel decisions. Abrami *et al.* (2007) did not address well-documented affective influences on such global ratings (Ambady and Rosenthal, 1993; Hamermesh and Parker, 2004). These influences show even stronger correlations between ratings and affect than between ratings and learning, although both correlations are positive and statistically significant.

To try to get around the problem of global ratings items, the University of Minnesota recommended calculating a global rating as the mean of their six core items. Deriving a global item indirectly in this way, from items that capture multiple dimensions, likely does screen out some of the superfluous affective influences such as physical attractiveness of the instructor that influence direct global ratings. However, based on Feldman's work, it seems that rather than a simple mean, that the means calculated from each representative dimension can be weighted according to their established importance. This possibly may produce an improved global expression over mere averaging. However, offering such a convenient general number invites the kind of incomplete evaluation noted by Leamnson's quote above, unless some procedural rule compels review committees to produce a written rating based on review of other multiple sources of information before getting to see the global expression.

McKeachie (2007) cautions against the temptation to use numbers as a basis for deducing teaching ability: "...faculty members should not use student ratings to compare themselves with other teachers." He goes on to note: "Once numbers are assigned, faculty promotion committees begin to make comparisons between teachers and assume that if one number is larger than another, there is a real difference between the teachers to whom the numbers have been assigned. Faculty members whose students 'agree' that they are excellent teachers may find that they fall as 'below average' in comparisons with those whose students 'strongly agree' on that same item."

Examination of the Student Ratings Form, its dimensions and its items

This section describes the ratings form proposed, dimension-by-dimension and the items drafted to represent each dimension. Dimension names come from Feldman (1998, 2007) and the definitions of the dimension come from Michael Theall (2007, personal communication). Feldman (1998) also used regression analyses to rate the percentage of the variance in student achievement generally explained by the dimension. "Achievement," as noted above, was defined by scores on tests and grades.

Proposed items	Ranked importance to achievement	References
DIMENSION 1: Preparation and organization (explains 30-35% variance)-the degree to which course, instruction, materials appear to be structured and arranged so as to promote learning	**1**	Feldman, 1998, 2007; Abrami *et al.* 2007; Weimer, 1991
1. I understood the learning outcomes expected from this course.		
2. To me, the course content seemed well organized.		
3. To me, class sessions seemed well organized.		

For this dimension, and for the ratings form in general, we avoided writing teacher-centered items such as "The teacher was organized" because students cannot judge the organizational vision teachers hold in their minds. However, students are the ultimate authorities qualified to report on the degree to which they can perceive an order to the course material and their experience of learning it.

Although teachers have a responsibility to enact a plan to help students to see order and interconnections in reasoning, teachers do not have full control of the results. As Leamnson (1999) noted, both learning and the decision to learn take place entirely in the minds of the students. Students' ability and desire to learn vary and are unique to every individual. Nevertheless, instruction that students perceive as organized, on the whole, produces learning that is superior to instruction that students perceive as lacking in organization. This dimension belongs to the two most important dimensions related to student learning. It is beneficial to students' learning to convey to them an awareness of order. Thus, gathering some information on what students perceive about this order is worthwhile.

Preparation enables teachers to promote students' awareness of organization at the scales of both the course and individual class sessions. While students cannot directly observe or experience a teacher's preparation, they can perceive some actions and products that result from it.

At the course scale, consistence between catalog descriptions, course syllabi, supplementary materials, what teachers enact in courses and what teachers evaluate is important. Such consistence constitutes an aligned organization that produces reinforcement of learning. Students are generally aware of the degree of this consistence or its absence. If the learning outcomes a teacher wants to instill are not carefully conveyed and developed, it is unlikely that students in general will be able to focus on these sufficiently to internalize them.

At the scale of class sessions, organization is conveyed by the outlining of key points or concepts near the beginning of a class, summarizing these at the end of the class session, and, at times, relating these to the larger outcomes of course and the signature attributes of education.

In summary, organization might be thought of as the quality of the map that we provide for students to help them to best arrive at the destinations we want them to reach. Preparation is analogous to the effort that we expend in creating the best map we know how to make.

DIMENSION 2: Clarity and understandableness (explains 24-30% variance)- the teacher's ability to express ideas and concepts in ways that students can understand	**2**	Feldman, 1998, 2007; Abrami *et al.* 2007; Boex, 2000
4. The time I spent in class sessions furthered my understanding of the course material.		
5. Examples and illustrations provided in this course aided my understanding.		
6. The course provided some general concepts that helped me see connections among specific topics.		

If preparation and organization can be thought of as a map, then Dimension 2, clarity and "understandableness" (a term coined by researcher Kenneth A. Feldman), are analogous to the attributes of a guide leading the students on an adventure across an unfamiliar landscape. A good adventure should certainly be challenging, but one measure of the quality of a guide lies in getting those who rise to the challenge of travel safely to their destination. A good guide won't try to carry the travelers, but likewise won't go off on her/his own adventure during conditions when the travelers become unlikely to make progress to their destination without help.

Most undergraduates don't have well-honed self-assessment skills, but in general they know when they are learning and when they feel dissatisfied with their level of progress. They know this well enough to make their reactions to items that relate to clarity and understanding sufficiently consistent to make this dimension a valuable measure. Producing clarity and understanding always ranks first or second in importance in meta-analyses that look at which teaching traits produce the best learning.

When researchers examine clarity and understanding in lecture classes, most of the dissatisfaction of students with teachers relates to lack of instructor clarity that results in students' lack of understanding. For lecture-based classes, it may be reasonable to write items that are teacher-centered such as "The teacher provided examples and illustrations that aided my learning." With group learning, however, it is often students within the group who produce the critical analogy or illustration that leads to other students' understanding. Thus, it becomes especially clear why an institution that emphasizes group learning through interaction should have a ratings form that is pedagogically independent.

Research overwhelmingly shows that students teaching other students constitutes one of the most effective learning arrangements. However, it is not just a matter of throwing students into small groups to talk about a topic. Instead, it is the good *design* of the learning exercise enacted within small groups that accounts for the great increases in small group learning. Even if a student writes: "I learned little from the teacher but much from my fellow students in this class," the teacher should be credited because the teacher designed the groups and learning experiences that produced this laudable result. The learning-centered items allow both the lecturer and the designer of group learning to receive proper credit under this category.

Increasing clarity and understanding does produce increased learning. One way to increase understanding regardless of chosen pedagogy is to meet students where they are and initially engage a topic by letting students connect with it based upon whatever they do know. Another is to come at new material in several different ways so as to enlist more regions of the brain in engaging material. Learning at the neurological level involves building synaptic connections. The more regions of the brain that are involved by engaging content in varied ways, the more numerous the connections being established.

In general, interactive engagement in well-designed group work produces more synaptic connections than does simply paying attention to a good lecture. However, group work done poorly may produce even worse results than poor lectures. Work with a faculty developer is one of the best ways to learn quickly how to increase clarity and understanding.

DIMENSION 3: Perceived outcome or impact (explains 15-20% variance)- the course relevance and academic or other outcomes; opportunities for success in the course	**3**	Feldman, 1998, 2007; Abrami *et al.* 2007; Bransford, Brown, & Cocking, 2000, Hildebrand et al., 1971.
7. The course was a valuable learning experience for me.		
8. The assignments in this course aided my learning.		
9. I was able to effectively use instructor feedback to increase my learning.		
10. I learned ways of reasoning that I could apply to other disciplines		

Dimension 3 addresses student perceptions of the value of the course to their lives. Perceiving a benefit to learning plays a significant role in students' motivation to learn. Just as students can't be expected to have mastered understanding of learning of the substance of a course before they have had the course, they can't always be expected to grasp the relevance of that substance to their own lives without an instructor's help in making this awareness more explicit. The research is clear that these particular perceptions play a role in aiding students to succeed, so it is worth making efforts to convey such awareness and to obtain information on the degree to which these efforts seem to be succeeding.

At Channel Islands, this may be a more difficult challenge than elsewhere, because of our stated and enacted emphasis on interdisciplinary thinking. Elsewhere, a teacher may be

able to clarify relevance to students based on utility of the discipline. Here, we need to convey the relevance of the substance of our course *and* contribute to students' respecting and valuing interdisciplinary thinking and learning. Most CSUCI faculty were educated in a discipline, but interdisciplinary thinking is challenging even for those faculty who were hired through an interview process that screened for ability and desire to support an interdisciplinary approach. It is understandably difficult to convey awareness for such expansive ways of thinking to students who as yet have little experience in thinking and learning within even a specialty area.

Practices within teachers' control can help students to perceive such relevance. Instructors can design assignments and grading rubrics that carry a reflective, metacognitive component to direct students to reflect on relevance. Likewise, our feedback can include statements that guide students to consider the learning in the light of relevance and value to themselves.

Proposed items	Ranked importance to achievement	References
DIMENSION 4: Stimulation of interest in content (Explains 10-15% of variance)- the inspiring and motivating of students to become engaged with content and course material; demonstrating relevance and importance	**4**	Feldman, 1998, 2007; Bain, 2004; Ryan & Deci, 2000
11. My learning experience increased my appreciation for the subject covered.		
12. I gained awareness of the relevance and importance of the course material.		
13. The course made a relevant contribution to my overall education.		

Dimension 4 is closely related to the preceding dimension, and methods used to promote appreciation for learning described under Dimension 3 can also advance appreciation in Dimension 4. Dimension 4 rates as the most important dimension to overall student satisfaction registered by global ratings, but it is subordinate to at least two others in generating student learning. Once again, the dimension has sufficient merit to learning to make it worth measuring.

Sometimes workers combine both dimensions under "motivation," (see Ryan and Deci, 2000) but there are benefits from considering these as the separate dimensions deduced by several workers and described by Feldman (2007). While Dimension 3 involves appreciation of relevance of learning to the person, Dimension 4 involves appreciation of learning brought about by increasingly valuing the knowledge for itself. To be sure, both dimensions are likely more strongly bound with affective qualities than cognitive thinking. However, no cognitive learning or decision to learn exists separately from affective feelings. Affective development is possible through learning, just as is cognitive learning.

Once an individual rejects learning through a decision of "I'm not interested in..." there is little chance that even the most intelligent mind can progress in gaining awareness of any learning it rejects on the basis of feelings. A teacher cannot hope to reverse such decisions of

every individual. However, student ratings are aggregate data of class averages, and the vast majority of students do not enter college with commitment to being intransigent.

Feldman (1986) deduced that, as teachers, our own self-confidence and our own enthusiasm for teaching are the two most important characteristics to our own success. Indeed, we convey these qualities to our classes and to our colleagues, whether or not we make deliberate efforts to do so. In older ratings forms designed largely for lecture classes, the teacher-centered item, "The teacher was enthusiastic" carried high association with the overall satisfaction outcomes. However, placing full responsibility on teachers to inspire their students to learn is an unfair and misguided use of ratings instruments. By writing ratings items that are learner-centered and focused on the learning, we can emphasize the nature of the classroom and the work being done.

Dimension 4 is too important to ignore as a component of ratings. However, many factors that influence a student's learning and enthusiasm for learning are unrelated to a teacher's inspiring of students. This is particularly true in interactive group learning, where other students' interest and relationships with course topics and substance can strongly influence students with lesser initial enthusiasm for learning that material.

DIMENSION 5: Encouragement and openness (Explains under 10% of variance)- the teacher's approachability ability to project and establish atmosphere that maintains motivation and promotes student - teacher interaction	**5.5**	Feldman, 1998, 2007; Abrami *et al.* 2007; Baxter-Magolda, 2001; Rheinberg et al., 2005
14. I felt I was evaluated fairly in this class.		
15. I felt I was treated with respect in this class		
16. The class atmosphere supported my learning.		
17. I felt encouraged to contribute civil dialogue to this class.		

By Dimension 5, we arrive at characteristics that exert measurable but lesser influence on learning in general. These qualities involve the concept of the "affective field" of a classroom (Rhem, 2008). The concept of "affective field" also extends beyond classes to departments and institutions (Nuhfer, 2008). The characteristics of a good affective classroom field involve encouragement of reasoned dialogue, trust, fairness, and respect.

There are many ways to encourage students in their quest for learning, and instructors provide encouragement in ways that vary with their experience. Most with experience have tried ways that proved effective and discovered others that backfired, especially when comments intended to use humor for encouragement became interpreted as sarcasm. Fear of stinging sarcasm can make students fearful of initiating interaction with their teacher.

Teachers are responsible for providing, within the constraints of what they can indeed control, a learning atmosphere in the classroom that is conducive to learning. Class disruptions can occur from a few individuals, and these require an instructor to act to maintain a good

learning environment. Skillful action will deal with the disruption in focused ways; an unskilled reaction may be perceived as a rebuke of the entire class.

It is important to use the syllabus to convey protocols that we believe are necessary for creating a class with a positive affective field. By sharing power and responsibility with students to help build trust and share enforcement of protocols, maintaining a healthy affective field becomes easier. People tend to take the best care of those things for which they feel some ownership.

DIMENSION 6: Availability and helpfulness (Explains under 10% of variance)- the personal availability and willingness to assist students; degree to which assistance results in improvement	**5.5**	Feldman, 1998, 2007; Baxter-Magolda, 2001; Palmer, 1998; Light, 2001
Rate the following three items only if you have first-hand experience from making contact for help.		
18. When I sought outside help from the instructor (such as by phone, e-mail or office visit), I received it.		
19. I felt welcome to seek help and advice from the instructor.		
20. The help I received from the instructor was useful to my learning.		

Dimension 6 ties with Dimension 5 in terms of its overall importance to learning. Older studies showed that, in large freshman classes at least, only a few percent of students ever sought out a professor in office hours, and perhaps this is a reason that Dimension 6, in general, does not rise higher in importance in meta-analyses. It examines a particular perspective not captured by peer visits to classes: how students and professors interact outside the classroom. Opportunity for such interactions varies with classes and institutions. The advent of e-mail makes professors more accessible to students on busy schedules, and this likely changed the dynamics of access reported by older studies.

Because CSU Channel Islands prides itself on placing students at the center and places an emphasis on student-teacher interaction, this is an important dimension for us to capture. One of the characteristics separating highly successful teachers from those less successful is the way that the former interact with students outside of class (Bain, 2004).

The items here look for the qualities of responsiveness when students do seek help rather than the numbers of times students seek help, which teachers cannot control. The cautionary notes on these items are needed, because in situations where only a small percentage of students seek individual help, a much larger group of students feeling compelled to enter something on the form will overwhelm the results from responses of the students who have sought a professor's help and actually know the answers to the items.

When students seek a teacher's help, it usually requires initiative and a small act of courage on their parts to admit that they need help. Sitting in lectures requires no interpersonal skill, but making a face-to-face visit to learn does. When students appear for an office

appointment, it is best to remain aware that they are more vulnerable than at other times, and their feeling welcome to seek this help and trusting that they will receive it are important. In instruction, teachers can help students to learn to self-assess in order to understand better when they truly need help and when they can best learn through their own efforts. When students seek help, assisting them up to the point where they can discover what they seek by taking that final important step for themselves in our presence is usually a good practice. Too much help that supplies all of the answers denies students such discovery. More help is not necessarily better (Svinicki, 2008).

Calculated Summary Item

Rather than solicit ratings from a direct global item that is prone to superfluous affective influences (Ambady and Rosenthal, 1993; Hamermesh and Parker, 2004), we opted in our draft version to calculate an average value for each dimension based upon the items under that dimension, and then producing a calculated summary item based approximately on Feldman's (1988) published contributions to variance in learning as follows

Dimension 1 mean x 0.30 = A
Dimension 2 mean x 0.30 = B
Dimension 3 mean x 0.15 = C
Dimension 4 mean x 0.15 = D
Dimension 5 mean x 0.05 = E
Dimension 6 mean x 0.05 = F

Calculated summary item = Sum A through F

As discussions continued in the task force, one person felt strongly that a summary item should be produced in another way. We realized that we did not need to make such a decision—that what, how and to whom results are reported is actually in the realm of policy and/or procedure, and our charge only was to produce a good form. As we pass this on to the Senate and the committee that will do this we are not recommending any particular summary in this final report. For information purposes only, we include in Appendix E some options we considered and insights we gained in our deliberations.

Open-ended items

No matter how much research goes into a student ratings form, none of the experts who did the research or drafted the form actually experience the course. Students who experience the course are in the best position to contribute suggestions. At the same time, open-ended items are a place in which a small percentage of students writing with anonymity can "vent" in ways that are abusive and totally inappropriate. In accord with McKeachie's suggestion that student ratings be a way in which to gather information in ways that help students to think metacognitively and accept responsibility, we devised three open-ended items through which to elicit additional information.

21-23. Knowing what you know now about the course, if it were possible to turn back time and you could experience this course again...

21. *What changes would you make in your own approach in order to improve your learning?*
22. *What aspects of the course would you advise your instructor to retain?*
23. *What suggestions would you provide to your instructor for revisions that would produce a better learning experience for you?*

Instructors often report that open-ended responses from students provide some of the most useful information to their making improvements in themselves and their courses. At the same time, opportunities to provide open-ended comments with anonymity constitute an opportunity for abuse. The wording in some forms' open-ended items almost invites abuses. In general, nothing should be said on a student ratings form that would not be said openly in civil conversation between adults. Irresponsible, abusive insults are devastating to professors' self-esteem and confidence. If permitted, the cumulative damage these do contributes to burnout, poorer service to future students, and erosion of the investment the institution makes in its faculty. By framing these questions in the context of having a face-to-face conversation with the professor, we hope that this will remind more students to make responsible statements and offer criticism in ways that are useful.

The updated proposed student ratings form for CSUCI, now titled "**Student Feedback on Course Learning Experiences**" appears in Appendix C. This version includes directions for completion with a generic bubble sheet form now being used as faculty volunteers pilot this form in their own classes now in spring, 2008. Slight wording format changes will be needed to deliver the form in other venues.

The updates result from comments and suggestions received by faculty from four "open house sessions," on the ratings form, emails by faculty to chair and members of this task force, and drop-in visits to the task force chair's office. All input received was shared with the entire task force. All modified items were checked to verify they passed the twenty-five screening criteria.

Some concern was voiced on several items that call for an affective response. The affective wording was retained in some items, and research justification for such wording was added in Appendix D.

References Cited

Abrami, P. C., d'Apollonia, S., and Rosenfield, S. (2007). The dimensionality of student ratings and instruction: What we know and what we do not. in R.P. Perry and J.C. Smart, (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 385-546.

Abrami, P. C., Rosenfield, S. and Dedic, H. (2007). The dimensionality of student ratings and instruction: An update on we know, do not know, and what we need to do. in R.P. Perry and J.C. Smart, (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 446-456.

Ambady, N. and Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. Journal of Personality and Social Psychology. *64*, 431-441.

Arreola, R. (2006). *Designing a comprehensive faculty evaluation system* (3rd ed). Bolton, MA: Anker Publishing.

Bain, K. (2004). *What the best college teachers do.* Cambridge, MA: Harvard University Press.

Barr, R., & Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. Chang., (November/December) 13-25.

Baxter-Magolda, M. (2001). *Making their own way: Narratives for transforming higher education to promote self-development.* Sterling, VA: Stylus Publishing, LLC.

Berk, R. A. (2006). *Thirteen Strategies to Measure College Teaching*. Sterling, VA: Stylus.

Boex, L. (2000). Identifying the attributes of effective economics instructors: An analysis of student evaluations. *Journal of Economic Education, 31*(3), 211-227.

Bransford, J., Brown, A., & Cocking, R. (Eds.) (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Cashin, W. E. (1988). Student ratings of teaching: a summary of the research: Kansas State Univ. Center for Faculty Evaluation and Development, Idea Paper n. 20. Available March 31, 2008 from http://www.idea.ksu.edu/resources/Papers.html.

Eder, D. (November, 2006). Assessment vs. evaluation: Dealing with the differences. Paper presented at the 2006 Assessment Institute, Indianapolis, IN.

Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: a review and synthesis: Research in Higher Education. *24* 129-213.

Feldman, K. A., (1998). Identifying exemplary teachers and teaching: evidence from student ratings: in *Teaching and Learning in the College Classroom* 2nd ed., K. A. Feldman and M. B. Paulsen, (Eds.), Needham Heights, MA: Simon & Schuster. 391-414.

Feldman, K. A. (2007), Identifying exemplary teachers and teaching: evidence from student ratings. in R.P. Perry and J.C. Smart, (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 93-129.

Hamermesh, D. S., and Parker, A. M. (2004). Beauty in the classroom: Professorial pulchritude and putative pedagogical productivity. Economics of Education Review. 2005 (pdf version available on 7/11/04 through Hamermesh's University of TX personal web site accessible through http://www.utexas.edu/.

Hildebrand, M., Wilson, R. C., and Dienst, E. R. (1971). Evaluating university teaching: Handbook published by Center for Research and Development in Higher Education, U of CA, Berkeley.

Hines, C. V., Cruickshank, D. R., & Kennedy, J. J. (1985). Teacher clarity and its relationship to student achievement and satisfaction. American Educational Research Journal, 22, 87-99.

Joint Committee: The California State University, California Faculty Association and Academic Senate CSU, (2008). Report on Student Evaluations of Teaching. Long Beach, CA: 13 p.

Langley, D., and others. (2007). Recommendations on modifying the student evaluation of teaching form at the University of Minnesota. Final Report of the Ad Hoc SCEP/SCFA Subcommittee, March 1, 2007, Revised report November 8, 2007.

Leamnson, R., (1999). T*hinking About Teaching and Learning - Developing Habits of Learning With First Year College and University Students*: Sterling, VA: Stylus.

Light, R. (2001). *Making the most of college: Students speak their minds.* Cambridge, MA: Harvard University Press.

McKeachie, W. J. (2007). Good teaching makes a difference—and we know what it is. in R.P. Perry and J.C. Smart, (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 457-474.

Nuhfer, E. B. (2008).The affective domain and the formation of the generator: Educating in fractal patterns XXIII part 1: National Teaching and Learning Forum, *18/2* 8-11. and part 2, National Teaching and Learning Forum, *18/3* 9-11. Available from any computer on the CSUCI campus at http://www.ntlf.com/restricted/.

Nuhfer, E. B. and Dewar, J., (2008). "Guidelines for Creating Good Items for Student Ratings of Professors." SoCal Faculty Developers Learning Community. Los Angeles, CA: http://www.lmu.edu/pagefactory.aspx?PageID=41603.

Nuhfer, E. B., (2007). A fractal thinker looks at student evaluations. An updated version of paper delivered at the 2005 AERA Meeting in Montreal in Theall, M., Abrami, P.C., Arreola, R., Franklin, J., Nuhfer, E., and Scriven, M. (2005). Valid faculty evaluation data: Are there any? AERA Annual Meetings Program Interactive Panel Presentation, American Educational Research Association Symposium, Montreal: April 14, 2005. 240. (summaries available at http://www.cedanet.com/meta/AERA2005valid.pdf and http://profcamp.tripod.com/fractalevals07.pdf)

Pascarella, E. T., Seifert, T. A., and Whitt, E. J., (2008). The role of classroom instruction in first-year persistence in college: Some new evidence. New Directions in Teaching and Learning, Jossey-Bass, in press.

Rheinberg, F., Vollmeyer, R., and Rollett, W., (2005). Motivation and action in self-regulated learning. in Handbook of Self-Regulation. Monique Boekarts, Paul Pintrich and Moshe Zeidner (eds.) Elsevier, 503-529.

Rhem, J., (2008) The affective field. National Teaching and Learning Forum, *18/3* 4-5.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*, 54-67.

Schonwetter, D., Menec, V., & Perry, R. (1995). An empirical comparison of two effective college teaching behaviors: Expressiveness and organization. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA: April, 1995.

Svinicki, M. (2008). When does enough feedback become too much? National Teaching and Learning Forum, *17/3* 12.

Theall, M., and Feldman, K. A., (2007) Commentary and update of Feldman's (1997) "Identifying exemplary teachers and teaching: evidence from student ratings." in R.P. Perry and J.C. Smart, (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 130-143.

Weimer, M. (1991). *Improving college teaching: Strategies for developing instructional effectiveness.* San Francisco: Jossey-Bass.

APPENDIX A
DRAFT: Minimal Requirements for an Online Ratings Delivery System for Student Ratings

Prepared by Edward Nuhfer, Director of Faculty Development, CSUCI

(1) The system should be centrally managed and remain compliant with campus policies. The system also has to be secure, so that it (a) cannot be entered by unauthorized persons who want to snoop through faculty ratings, students' comments, or change any of it; (b) that it prevents individuals from abusing the system by providing more than one rating per course section, and (c) that it logs access and requires identification of anyone who enters the system at anytime. That includes the manager of the system. Any vendor should be able to provide a track record of security.

(2) The system must accept our CSUCI Survey and algorithms that we choose to use to produce numerical ratings.

(3) The system must provide access to multiple blocks of items that include (a) the core items selected by CSUCI for evaluative personnel decisions; (b) an optional formative survey that can be selected for improvement; (c) an customizable option wherein related units such as a department, a school, a center can have a block of items added that are specific to courses under their area of responsibility and (d) a customizable option for the individual faculty member to add an item or block of items that the individual is interested in tracking.

(4) The system must provide multiple levels of access for selective distribution of survey response results, such that the essential information from specific survey items is distributed to those in accord with their specific evaluative responsibilities. For example, data gathered from items related to a specific school or center would not go to the RTP review committee; the RTP committee has no responsibility for running a specific school or center. Results from items created by the instructor would return only to the instructor.

(5) The system must permit text entry by respondents to open-ended questions as well as graded response settings.

(6) The system must integrate with our CSUCI course enrollment system so as to allow only students enrolled in a specific course to provide ratings on that course and on the instructor responsible for that course.

(7) The system should permit useful and easily customizable reporting of the data.

(8) There should be manual over-rides for any automated system so that special courses such as team-taught courses or course taken over in mid-semester by a second instructor can be pulled from the automated processing and handled in ways that make sense.

(9) Data must be secured in a safe place so that loss by "human error" or other mishaps is at an acceptably low risk. Provisions should allow data and reports to be archived indefinitely.

(10). The system should come from a vendor who provides training, documentation, and maintains responsive phone and web support.

(11) The system should be cost-effective relative to paper scoring sheet systems.

(12) The system must meet the accessibility standards required by law.

Additional desirable qualities are ability of the system to identify and remind students by email who have not completed their ratings and ability of the system to be used for special purposes beyond student ratings such as surveys or online campus voting, and ability to export data into other tools such as Excel, or statistical and graphing software, for custom reporting and analyses.

APPENDIX B

PART 1. Screening Criteria Applied to Each Item on the Student Ratings Form

(Querying criteria based on standards from Berk, 2006; Nuhfer and Dewar, 2008, and drafted to design a match between the Student Ratings Form and the Mission Statement of the California State University Channel Islands <www.csuci.edu/about/mission.htm>

The item:

1. Avoids favoring a particular discipline type?
2. Is pedagogically independent?
3. Is primarily learner-centered?
4. Is clear and direct?
5. Is brief?
6. Addresses one topic?
7. Is a sentence?
8. Is written at an appropriate reading level?
9. Is grammatically correct?
10. Is strongly worded?
11. Fits within its intended dimension?
12. Describes a positive or negative?
13. Applies generally to all?
14. Is a statement that the respondent is qualified to answer?
15. Avoids having a double meaning?
16. Is devoid of double negatives?
17. Is not absolute fact?
18. Is not "fuzzy" (has no non-absolutes like "generally" or "often.")?
19. Is not inflammatory or value-laden?
20. Is devoid of jargon?
21. Is unique-not duplicative of other items?
22. Is scalable- not merely a factual statement with a unique right answer?
23. Is free of endorsement-doesn't request an answer most should already know from a common source such as the syllabus)?
24. Addresses what the item intended to address?
25. Describes something that the teacher can control or influence for the class as a whole. Faculty development in support of this is possible.

Part 2. Some Guidelines for Creating Good Items for Student Ratings of Professors

(Compiled by Ed Nuhfer and Jackie Dewar. Examples of flawed items contributed by members of POD. Special thanks go to Mike Theall and David Langley for advice. Chapter 3 of Ronald Berk's *Thirteen Strategies to Measure College Teaching*, 2006, is highly recommended for writing rating items.)

Students can offer their authority of experience in

1. Teacher/student relationships experienced by personal performance (including rapport, respect for students, *etc*.)
2. Experience of a supportive classroom environment and the management of that environment
3. Instructional practices they experienced (*e.g*. organized presentation, feedback, clarity)
4. Issues about the level of student engagement (*e.g.*, amount of effort expended, time spent on studying, workload, and difficulty)
5. Congruence between evaluation measures (tests) and what was taught
6. Accessibility to the instructor
7. Extent to which the syllabus served as a useful guide for the course...

...if the items to elicit their experiences are well constructed.

Faculty have authority of expertise in

1. Content and what is appropriate and sufficiently current
2. Relationship of the course to the overall curriculum or program
3. What they think and feel vs. what others say they think and feel
4. Course design issues such as
    A. Appropriateness of course objectives & outcomes
    B. Appropriateness of teaching methodology in relation to content
    C. Overall presentation sequence and conceptual framework of the course

...so, avoid inviting students to second-guess the faculty expertise.

EXAMPLES

1.) Don't ask students to judge characteristics beyond their experience/expertise.

Flawed Items

a. The teacher was knowledgeable.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. This teacher respects students.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. The course objectives were appropriate.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

2.) Measure only single constructs with single brief items.

Flawed Items

a. To what extent did this course advance ability to read with comprehension, and to write and speak clearly, coherently, and effectively as well as to adapt modes of communication appropriate to an audience?
(very much some not much not at all not applicable)

b. The instructor is punctual in meeting class and office hour responsibilities.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

c. Course activities (discussions, small group work, labs, projects, lectures, exams, etc.) were clearly prepared, organized, and sequenced to help students achieve course goals.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

3. Avoid items that can lead to ratings without regard to context.

Flawed Items

a. If your schedule allowed it, would you take another course taught by this professor, or would you recommend the professor to someone else?

(The possible responses are Yes/No with a space to elaborate.)

b. As a result of this course, I began to challenge the opinions of others.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

4. Avoid items that rate professors based on circumstances they cannot control.

Flawed Items

a. I liked the time this class was offered. (Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. The classroom was comfortable. (Strongly agree 5) 4 3 2 (1 Strongly disagree)

5. Avoid printing poorly crafted items that make little sense.

Flawed Items

a. The teacher encourage students to ask ideas.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. The course was sensitive to diversity
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

6. Within reason, students should be able to read and understand any item.

Flawed Items

a. The teacher taught at the appropriate Perry stages for this class.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. The professor's paradigm was parsimonious grading.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

7. Items should solicit first-hand information experienced by the respondent.

Flawed Items

a. The teacher was available for help after 5:00 a.m.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. Students believe this class is easy.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

8. Items should avoid both nonspecific and absolute terms.

Flawed Items

a. The instructor usually inspired me to attend class.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

b. Every class began promptly on time.
(Strongly agree 5) 4 3 2 (1 Strongly disagree)

9. Avoid items that can have double meanings.

Flawed Items

a. I respect this professor. (Strongly agree 5) 4 3 2 (1 Strongly disagree)

10. Scrutinize any item that does not map to an established dimension of instruction, and plan to use all information you gather for the purpose intended. (Many of the above items fail on usefulness alone. Don't mix survey items about unrelated topics into the faculty evaluation items.)

APPENDIX C: Student Feedback on Course Learning Experiences

Leave name, ID, and all personal identification information blank on the Parscore® Form. Please use the following scale for rating each item.

Blacken the **A** bubble**—Strongly Agree**
Blacken the **B** bubble—**Agree**
Blacken the **C** bubble—**Disagree**
Blacken the **D** bubble—**Strongly Disagree**
Blacken the **E** bubble—**Not Rated/Not Applicable**

1. I understood the learning outcomes expected from this course.
2. To me, the course content seemed well organized.
3. To me, class sessions seemed well organized.

4. The time I spent in class sessions furthered my understanding of the course material.
5. Examples and illustrations provided in this course aided my understanding.
6. The course provided some general concepts that helped me see connections among specific topics.

7. The course was a valuable learning experience for me.
8. The assignments in this course aided my learning.
9. I was able to effectively use instructor feedback to increase my learning.
10. I learned ways of reasoning that I could apply to other disciplines.

11. My learning experience increased my appreciation for the subject covered.
12. I gained awareness of the relevance and importance of the course material.
13. The course made a relevant contribution to my overall education.

14. I felt I was evaluated fairly in this class.
15. I felt I was treated with respect in this class
16. The class atmosphere supported my learning.
17. I felt encouraged to contribute civil dialogue to this class.

Rate the following three items only if you have first-hand experience from making contact for help.
18. When I sought outside help from the instructor (such as by phone, e-mail or office visit), I received it.
19. I felt welcome to seek help and advice from the instructor.
20. The help I received from the instructor was useful to my learning.

Open-Ended Items – Please Respond on back of ParSCORE® form

21-23. Knowing what you know now about the course, if it were possible to turn back time and you could experience this course again...

21. *What changes would you make in your own approach in order to improve your learning?*
22. *What aspects of the course would you advise your instructor to retain?*
23. *What suggestions would you provide to your instructor for revisions that would produce a better learning experience for you?*

APPENDIX D
Reasons for Affective Wording of Some Student Ratings Form Items

In brief, the affective domain is inseparable from the cognitive in teaching and learning. How people feel about learning affects how well they learn. How people feel about their working and learning environment, at least over the long term, affects their performance. CSUCI's chosen mission of "...integrative approaches, emphasizes experiential and service learning..." is in part affective by virtue of its being selected as a preference among other possible legitimate choices. Further, this stated type of learning experience provided by CSUCI seems designed to provide for growth in social awareness, concern for others, and affective development in concert with cognitive development. Research confirms that student responses to items calling for cognitive information are nevertheless responses that are largely determined by affective feelings. APPENDIX D carries the research justification for affective wording on some items examined in student ratings.

Educators and psychologists recognized three major domains of the mind: cognitive (Bloom and Krathwohl, 1956), affective (Krathwohl, Bloom, and Masia, 1964) and psychomotor (Dave, 1967; Harrow, 1972; Simpson, 1972). These volumes arose because all three domains are involved in learning and becoming educated. The affective domain volume proved to be decades ahead of its time.

After 1956, the taxonomy of the cognitive domain was immediately accepted. It remains cited often, perhaps more often than any single publication in higher education. Most professors have heard of "Bloom's Taxonomy" and many employ for designing detailed evaluation, assessment, goal-setting and class discussions. In contrast, only a few professors know of the other taxonomies' existence.

Perry's *Forms of Ethical and Intellectual Development in the College Years* (Perry, 1999) is the hallmark study on how individuals change through gain in cognitive abilities and awareness of affect. Inclusion of *Ethical* in his title formally tied Perry's model to including development of moral judgment and the affective domain as an outcome of higher education. Perry's lower six levels that mapped with Bloom's cognitive taxonomy were immediately embraced, whereas his reported upper levels 7, 8 and 9 were largely ignored until recently.

The importance of emotions had solid recognition long before Perry's study. History's greatest naturalist Charles Darwin (Darwin, 1872, 1890; Ekman, 1989) recognized it. William James, a recognized giant among psychologists recognized it (James, 1884; 1890), and Benjamin Bloom, one of the most influential educational researchers, was a major driving force behind the second volume on the affective domain noted above.

Yet, aversion to crediting importance to the affective domain pervaded all of higher education, and it influenced even brain scientists. Damasio (1999) notes that academicians' long-standing suspicions of the value of emotions provided a stigma that discouraged even psychologists from taking up the affective domain as a respectable topic of research (Damasio, 1999, p. 39). It's understandable that professors who are not current on studies of the importance of affect will still react dismissively toward items on ratings forms that solicit for affective feedback. Yet, all the learning dimensions are in part influenced by affect. On our proposed form, Dimensions 4, 5 and likely 6 are dominantly affective and still rise to the top half dozen dimensions of importance to cognitive learning.

Today, clinical studies and brain imaging techniques confirm that emotion is integral to the processes of learning, reasoning, and judgment (Damasio, 1999, Haidt, 2007; Reynolds,

2006; Tucker and others, 2003). Further, many noteworthy contributions to student success in college employ the affective domain as an asset in design for learning.

Kegan (1982) advocated a pronounced role for the affective domain in intellectual development. He ascribes an importance to the affective domain that gives it parity with the cognitive domain and places development of both on a parallel course.

> It should be clear that my conception neither subsumes affectivity to the cognitive realm, as traditional Piagetians tend to do...nor makes intellectual life the offspring servant of affect, as psychoanalysis tends to do. (p. 83). In arguing for evolutionary activity as the fundamental ground in personality, constructive-developmental theory is not choosing between "affect" or "cognition" as the master of development...but is putting forth a candidate for a ground of consideration prior to and generative of cognition *and* affect. (p.81)

Kegan's work represented an early shift in awareness, and it unabashedly makes abundant use of terms that describe affective qualities. LeDoux (1996), goes even farther, and vests the affective domain with the power to overrule the cognitive domain in the determination of important actions and decisions.

Albert Bandura (1997) contributed a lifetime to understanding the nature and importance of *self-efficacy*—a quality that contains one's affect and beliefs about her/his capabilities to succeed. The importance of self-efficacy is perhaps shown best by a brief quote from Mohatma Gandhi on the self-efficacy web page <http://www.des.emory.edu/mfp/self-efficacy.html>: "If I have the belief that I can do it, I shall surely acquire the capacity to do it, even if I may not have it at the beginning."

Self Assessment (Alverno College Faculty, 2000) is the basis for Alverno College's unique curriculum. It is the result of design work by Alverno's faculty, begun in the late 1960s. Alverno represents a departure from theoretical and longitudinal studies about what happens to students and moves to application in an unprecedented institutional commitment to increase students' competency in both the cognitive and affective domains. Development under Alverno's Self Assessment model includes probing of one's own "intuition, emotion and attitude" (Alverno College Faculty, 2000 p. 4) along with content learning in order to develop skills of self-efficacy (Bandura, 1997) and judgment. The last accords with Bloom's (1956) Level 6 evaluation, which when done well with affective awareness, places the student at the higher Perry (199) stages.

Baxter Magolda (2001) borrows heavily from Kegan (1982) for her model of "*Self authorship*" that presents four ways of knowing: absolute knowing, transitional knowing, independent knowing, and contextual knowing. Her research method is similar to that of Perry—longitudinal studies based on careful analyses of interviews. Her descriptions of ways of knowing are similar to the descriptions given by Perry for his developmental stages—sufficiently to permit these to be mapped with Perry's stages as well as onto a number of others' taxonomies of adult thinking.

Marzano (2001) reviewed Bloom's 1956 cognitive taxonomy, and Marzano and Kendall (2007) offer a replacement taxonomy —not to be confused with the updated ...*Revision of Bloom's Taxonomy of Educational Objectives* (Anderson and others, 2000). Marzano's model consists of a base of knowledge plus three systems that are not hierarchical: The Cognitive System, the Metacognitive System, and the *Self-System*. Recognizing one's own emotional influence in cognitive decisions is necessary to achieve a developed Self- System

(Marzano, 2001, p. 103) Marzano devotes pages to the place of emotion in its influence on motivation, which he notes influences academic success.

Self-regulation or *self-regulated learning* places emphasis on intrinsic motivation and models a good approach to learning as a cycle of forethought, performance control, and self-reflection (Zimmerman and Schunk, 1989). can produce this familiarity, and it should be in every developer's library. Attention to the affective domain is essential to the self-regulation model (Boekaerts, Pintrich and Zeidner, 2005; Ilies and Judge, 2005; Diamond and Aspinwall, 2003.) There is ample overlap with the reflective component of Alverno's self assessment and with Bandura's self efficacy to indicate that these approaches accord well with one another and with self-regulation theory.

Thus based on the nature of learning, the CSUCI mission and the research on student ratings, some queries that ask for affective responses such as feelings about specific situations and actions is fully justified.

References

Alverno College Faculty. (2000). *Self Assessment at Alverno College*. G. Loacker, (Ed.) Milwaukee, WI: Alverno College.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., Wittrock, M. C. (2000). *Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Boston: Allyn & Bacon.

Baxter Magolda, M. B. (2001). Making their own way: Narratives for transforming higher education to promote self-authorship. Sterling, VA: Stylus Publishing.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: Freeman.

Bloom, B.S. (1956). *Taxonomy of Educational Objectives— The Classification of Educational Goals*. *Handbook I. - Cognitive Domain*. NY: David McKay.

Boekaerts, M., Pintrich, Paul R. and Zeidner, M. (Eds.). (2005). Handbook of Self-Regulation. San Diego: Academic Press.

Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt.

Darwin, C. R. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.

Darwin, C. R. (1890). *The Expression of the Emotions in Man and Animals*. 2nd edition. Edited by Francis Darwin. London: John Murray.

Dave, R. (1967). Psychomotor domain. Berlin: International Conference of Educational Testing.

Diamond, L.M. and Aspinwall, L.G. (2003). Emotion regulation across the life span: An integrative perspective emphasizing self-regulation, positive affect, and dyadic processes. Motivation and Emotion, 27 (2) 125-156

Ekman, Paul ed. and Darwin, C. R. (1989). *The Expression of the Emotions in Man and Animals*. 3rd edition. London: Harper Collins.

Haidt, J. (2007).The New Synthesis in Moral Psychology. Science 316 (5827) 998-1002. May 2007: 998-1002

Harrow, A. (1972). *A Taxonomy of the Psychomotor Domain. A Guide for Developing Behavioral Objectives.* New York: McKay.

Ilies, R., and Judge, T.A. (2005). Goal regulation across time: The effects of feedback and affect. Journal of Applied Psychology. 2005, 90 (3). 453–467.

James, W. (1884). What is an Emotion? *Mind*, 9, 188-205. http://psychclassics.yorku.ca/James/emotion.htm

James, W. (1890) *The Principles of Psychology*. Chapter 10, Consciousness of Self, and Chapter 25, The Emotions. Available at <http://psychclassics.yorku.ca/James/Principles/prin25.htm>

Marzano, R.J. (2001). *Designing a New Taxonomy of Educational Objectives*. Thousand Oaks, CA: Corwin Press.

Marzano, R.J. and Kendall, J.S. (2007). *The New Taxonomy of Educational Objectives, 2nd edition*. Thousand Oaks, CA: Corwin Press.

Perry, W. G. Jr. (1999). *Forms of Ethical and Intellectual Development in the College Years*. San Francisco, CA: Jossey-Bass (a reprint of the original 1968 work with minor updating).

Reynolds, S. J. (2006). A neurocognitive model of the ethical decision-making process: Implications for study and practice. Journal of Applied Psychology, 91(4), 737-748.

Simpson, E. (1972). The classification of educational objectives in the psychomotor domain: The psychomotor domain. Vol. 3. Washington, DC: Gryphon House.

Tucker, D. M., Luu, P., Desmond, R. E., Hartry-Speiser, A., Davey, C., & Flaisch, T. (2003). Corticolimbic mechanisms in emotional decisions. Emotion, 3(2), 127-149.

Zimmerman, B J. and Schunk, D.H., (1989). *Self-Regulated Learning and Academic Achievement: Theory, Research, and Practice.* New York, Springer. Springer Series in Cognitive Development.

APPENDIX E
CONSIDERATIONS FOR SUMMARY REPORTING of DATA

The task force that created the "**Student Feedback on Course Learning Experiences"** did not attempt to set any policy about reporting. However, discussions we held may prove useful to those who set policy. Here are a few brief considerations.

1. Responses to open-ended items 21, 22, 23. Should these be given a numerical point value, used for qualitative interpretations of short-answer survey items, or simply passed on to the instructor?

2. Reporting of data on items 1-20. At minimum we should capture, means, standard deviations, number of valid responses to each item, and total respondents participating in the class rating. To whom should that detailed rating be transmitted? Benefits are details; drawbacks are "too much information," depending on the end user.

3. Summary option 1. The logic of the "Student Feedback..." form was based upon prioritization established in literature of teaching dimensions. A summary report can capture the information by calculating a rating for each dimension from the items designed to map into that dimension as follows.

Instructional Dimension	Numerical algorithm to express the summary score on the dimension	Ranked Importance with Learning Achievement
Preparation and organization	= Average of items1 to 3	1
Clarity and understandableness	= Average of items 4 to 6	2
Perceived outcome or impact	= Average of items 7 to 10	3
Stimulation of interest in content	= Average of items11 to 13	4
Encouragement and openness	= Average of items 14 to 17	5.5
Availability and helpfulness	= Average of items18 to 20	5.5

It makes sense that all averages be reported next to the correctly labeled dimension and the ranked importance of that dimension to learning achievement always carried with the average. Our logic of our form was to try to capture each of these dimensions with a few questions. Because we are trying to understand a fractal neural network associated with each dimension, a single measure (item) for each dimension is certainly inappropriate. We tried to create three items per dimension to achieve multiple measures of it. Two of the dimensions have four items. This is not a problem to the philosophy behind the form. The score of the dimension from the items bundled into it is what we are after and not the individual item scores. Benefits: reporting all six dimensions back to a faculty member provides a convenient way to target and choose specific areas to work on for improvement or to seek out assistance on from the faculty development director. It provides a convenient profile with a lot of information in a short table. Drawbacks are that in reality the neural networks involved in each of these dimensions are in constant communication with one another; these measures are not reflecting entities isolated from one another, even though their separate tabulations in a table may be erroneously interpreted as measures of characteristics that are cleanly separated. They are not. Another potential drawback noted by a faculty reviewer is that a single dimension might be seized upon by an RTP committee and be misused as THE student rating unless policy guidelines prevent doing this.

4. Summary Option 2. Produce a single summary item based upon weighting the scores on the six dimensions in accord with their general relative importance established in the meta-analyses as follows

Instructional Dimension	Numerical algorithm to express the summary score on the dimension	Ranked Importance with Learning Achievement
#1 Preparation and organization	= Average of items1 to 3	1
#2 Clarity and understandableness	= Average of items 4 to 6	2
#3 Perceived outcome or impact	= Average of items 7 to 10	3
#4 Stimulation of interest in content	= Average of items11 to 13	4
#5 Encouragement and openness	= Average of items 14 to 17	5.5
#6 Availability and helpfulness	= Average of items18 to 20	5.5

Dimension #1 summary score x 0.30 = A
Dimension #2 summary score x 0.30 = B
Dimension #3 summary score x 0.15 = C
Dimension #4 summary score x 0.15 = D
Dimension #5 summary score x 0.05 = E
Dimension #6 summary score x 0.05 = F
Calculated summary item = Sum A through F

The task force anticipated that an RTP committee would likely try to produce their own summary item by averaging items or averaging dimensions. Because meta-analyses shows that all items are not equally important, the task force thought that producing a weighted summary item based on relative importance of the dimensions established in the literature would be more appropriate than a simple average. At least one seminal paper on student ratings supported that view. However, one task force member from the mathematics department objected vehemently to this weighted average. The suggested alternatives this member provided are in the Summary Options below.

5. Summary Option 3—Simply average all twenty items. Although this may at first glance seem to violate the priorities established through research with regard to relative importance, it is important to recognize that the above dimensions were established largely on lecture-based classes. By creating a pedagogy-independent form, we are breaking new ground. For this reason it appears wise to report both the weighted option # 2 above and the simple average together, so that these can be quickly compared. It may be that personal attention in small active learning classes will make some dimensions of greater relative importance than indicated by the prevalent literature.

We are happy to let the policy committee choose a procedure for calculating and reporting a single summary item. It is probably needed, because review committees will likely try to do this anyway (unless a policy guidelines forbids this). Benefits: produces a summary that takes into account all of the information collected on every short item in the survey; offers a convenient summary to review committees. Drawbacks: The relative weightings come from meta-analyses, which clearly indicated a hierarchy of importance of dimensions. However, the specific weightings derived are not necessarily those that would

derive from a single form or small population of study. Finally, characterizing any faculty member by a single number, even a useful number, is a disservice, and the convenience of having a single number invites reviews that do not fairly consider all evidence that faculty members are required to supply in a portfolio. Taking this summary number as a way to rank faculty competitively against one another by putting names and numbers into a spreadsheet and hitting the "sort" button is not evaluation; it's abuse. Only policy can prevent this kind of misuse of the data.

Final noted policy gap.

Of more importance than the relative weightings and expressions on this form is the weight of student ratings as the part of the overall evaluation process. Faculty we spoke with expressed fear that the ratings form—or worse a single number derived from any form—would end up being THE evaluation.

Although policy states that other measures need to be taken into account, this is not sufficient if the policy allows THE rating to derive 99% from the ratings form and 1% from all other evidence submitted. Faculty and administration together should develop thoughtful policy guidelines that do not prevent an RTP committee from having reasonable flexibility, but offer sufficient clarity to let faculty know on what basis they are truly being evaluated. The fear noted above can be removed by improving clarity of the evaluation process.

Appendix F – Preliminary Results from Spring 2008

Faculty hosting twelve separate classes (containing total of about 250 students) volunteered to use CSUCI's newly approved "Student Feedback on Course Learning Experiences." The initial critical test involved a test for internal reliability. The tool proved to be as stable as the long-established larger commercial tools. The split-halves correlation using odd and even items (ten each) produced an r-value of 0.994, with a 95% confidence level of being greater than 0.97; an r-squared of 0.989, and a Pearson-Brown Reliability Coefficient of 0.997. The slope is 0.987 with an intercept of 0.007.

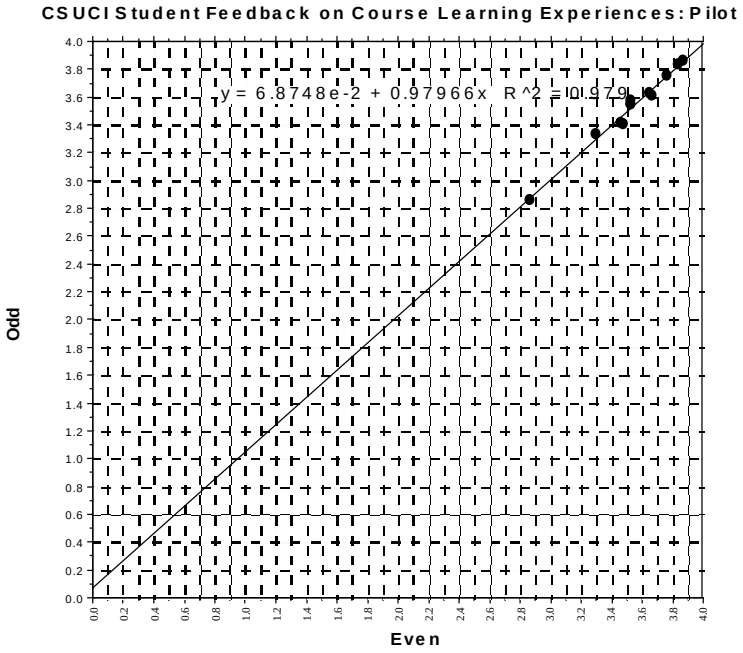

Figure 1. Least-squared line fit of twelve classes and statistical parameters.

The Director of Faculty Development reviewed all students' written comments supplied in response to the open-ended items 21-23. The comments were reflective and thoughtful. Further, they validated patterns indicated by the short open-ended items. Not one comment in the lot was of the "fire this professor" category or a comment that could not be said in a civil conversation.

Three of the professors who volunteered ran the survey in multiple sections of the same course, and one of these ran the survey in two different courses. This allowed us an initial look at consistency between same-course same-professor and different course same professor. The results were highly consistent (Figure 2 is typical) in both cases. The learning created as expressed by students' responses on the twenty items seems strongly related to instructors' practices.

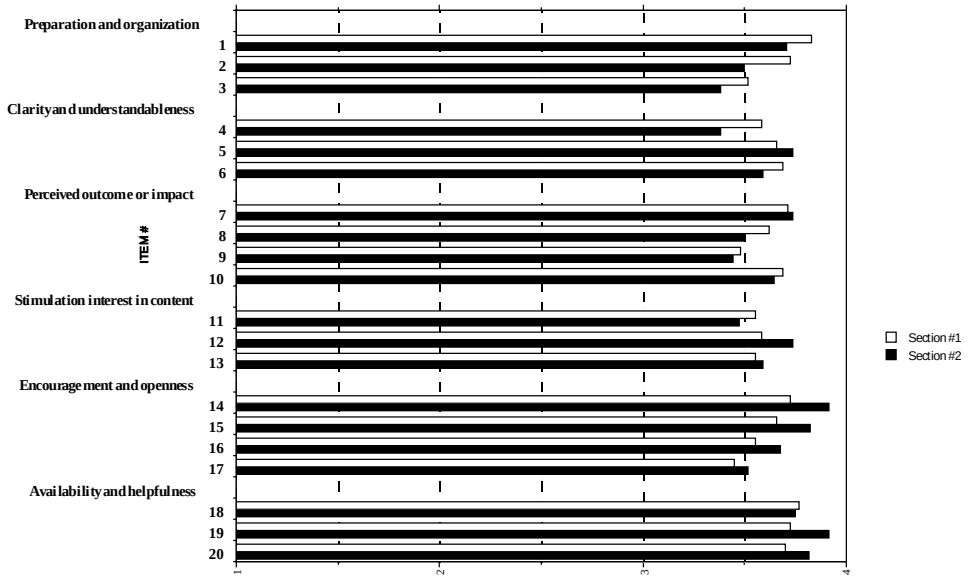

Figure 2. Results from Student Feedback on Course Learning Experiences: Same instructor; two sections of same course.

The director ran a factor analysis on the twenty items based on the pilot sample. The results appear in Figure 3. About 55% of the variance is explained by loadings of all twenty into Factor 1. Factor 2 accounts for 9% of the variance and is heavily loaded by Dimension 6 items. That item is contributed by out-of-class experiences of 161 students with faculty. Remember that students are asked on the form NOT to respond if they have no first-hand experience with asking a faculty member for outside help. This meant that about 161 out of 243 respondents (2/3) made use of an instructor's personal outside help. This first factor pattern matrix is shown in Figure 3.

Eigenvalues

	Magnitude	Variance Prop.
Value 1	11.066	.553
Value 2	1.793	.090
Value 3	.894	.045
Value 4	.747	.037
Value 5	.684	.034
Value 6	.613	.031

Unrotated Factors

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Item 1	.744	.175	.027	.363	-.097	-.014
Item 2	.794	-.078	-.196	-.186	-.369	.116
Item 3	.733	-.281	-.278	-.363	-.148	.043
Item 4	.761	-.191	.064	-.086	.430	.010
Item 5	.763	-3.490E-3	.157	-.150	.201	.261
Item 6	.772	-.193	.017	.256	.203	.118
Item 7	.859	-.163	-.126	.114	.127	.016
Item 8	.734	-.342	-.254	-.097	-.020	-.144
Item 9	.773	.064	.034	-.278	.131	-.049
Item 10	.773	-.213	.062	.164	-.067	-.304
Item 11	.806	-.193	-.075	.030	-.175	-.036
Item 12	.791	-.232	-.029	.340	-.071	.097
Item 13	.862	-.265	-.073	.038	.048	-.140
Item 14	.706	.242	-.061	.125	-.061	.483
Item15	.667	.332	.304	-.034	-.205	.103
Item16	.721	.021	.482	-.067	-.085	-.029
Item 17	.699	.026	.466	-.122	-.132	-.211
Item 18	.514	.645	-.256	.194	-.042	-.260
Item 19	.628	.645	-.179	-.161	-.032	-4.853E-3
Item 20	.699	.497	-.094	-.100	.318	-.111

Figure 3. Nonrotated factor pattern matrix with six components extracted via principal components analysis for all twenty items on Student Feedback on Course Learning Experiences form. Number of cases = 161.

In order to obtain data from all students who responded by completing all other survey items, the factor analysis was run again on the just the seventeen items representing the first five dimensions. Results shown in Figure 4 reveal a similar loading pattern into the first factor, which accounts for 54% of the variation. The second factor, explaining 7% of variance, is cleanly defined by the four items (14-17) that constitute the measure of Dimension 5.

Eigenvalues

	Magnitude	Variance Prop.
Value 1	9.116	.536
Value 2	1.203	.071
Value 3	.837	.049
Value 4	.704	.041
Value 5	.675	.040
Value 6	.613	.036

Unrotated Factors

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Item 1	.705	.222	-.086	-.256	.010	.168
Item 2	.765	-.064	.412	-.109	-.239	-.188
Item 3	.707	-.290	.522	-.077	-.046	-.100
Item 4	.734	-.173	.087	.176	.192	.367
Item 5	.712	-.027	.156	.311	-.161	.351
Item 6	.748	-.125	-.138	-.032	.112	.225
Item 7	.843	-.212	-.195	.035	-.084	-.042
Item 8	.713	-.256	.036	-.160	.202	-.241
Item 9	.714	.218	.208	-.251	.176	.146
Item 10	.723	-.073	-.256	-.097	.376	-.181
Item 11	.778	-.224	-.117	6.457E-3	-.182	-.137
Item 12	.774	-.150	-.326	.059	-.196	-.071
Item 13	.839	-.235	-.168	.073	-.011	-.010
Item 14	.667	.354	-.139	-.389	-.253	.158
Item15	.626	.551	-.051	.106	-.224	-.177
Item16	.717	.273	2.895E-3	.453	3.848E-3	-.153
Item 17	.646	.473	.136	.125	.366	-.113

Figure 4. Nonrotated factor pattern matrix with six components extracted via principal components analysis for seventeen items on Student Feedback on Course Learning Experiences form, after the Dimension 6 for outside help from the professor have been stripped. Number of cases = 243.

How do such factor pattern matrices compare with those produced by forms that have been refined for reliability and validity through strict psychometric tuning? We can see one such pattern for the form, Students' Evaluation of Educational Quality (SEEQ). This is a form that is commercially available and has a long history. The factor pattern matrix shown in Figure 5 was originally published by Abrami and d'Appolonia in Journal of Educational Psychology (1991) and is still reproduced today (Abrami, d'Apollonia and Rosenfield, 2007, p. 421).

Table 5: Nonrotated Factor Pattern Matrix with Six Components Extracted via Principal Components Analysis[1]

SEEQ Item	Factor loadings on first six components					
	I	II	III	IV	V	VI
Course challenging	.893	.392	.017	−.109	.037	−.056
Learned something valuable	.873	.262	.021	−.148	.133	−.080
Increased subject interest	.868	.122	−.013	−.226	.153	−.031
Understood subject matter	.760	−.182	−.094	−.167	.201	−.142
Overall course rating	.940	.176	−.032	−.085	.017	−.120
Enthusiastic about teaching	.886	.017	.039	−.082	−.308	−.073
Dynamic and energetic	.875	.104	.045	−.133	−.324	−.156
Enhanced presentation with humor	.787	.009	.039	−.158	−.315	−.152
Teaching style held interest	.884	.075	.012	−.147	−.249	−.208
Overall instructor rating	.941	.058	−.037	−.024	−.181	−.105
Explanations clear	.868	−.057	−.180	−.076	−.072	−.121
Materials prepared and clear	.857	.065	−.300	.052	−.064	−.060
Objectives stated and pursued	.855	.130	−.255	.125	.056	−.094
Lectures facilitated note taking	.649	.153	−.485	.118	−.158	.043
Encouraged class discussions	.741	−.390	.389	−.233	.104	−.068
Students shared ideas/knowledge	.688	−.479	.394	−.226	.136	−.046
Encouraged questions and answers	.852	−.303	.237	−.108	.047	−.070
Encouraged expression of ideas	.780	−.414	.349	−.131	.084	−.016
Friendly towards students	.769	−.370	.256	.258	−.097	.117
Welcomed seeking help or advice	.756	−.310	.246	.393	−.122	.184
Interested in individual students	.817	−.277	.261	.303	−.089	.117
Accessible to individual students	.678	−.180	.138	.471	−.108	.298
Contrasted implications	.803	.010	−.205	−.156	−.016	.418
Gave background of ideas/concepts	.819	−.024	−.237	−.201	.033	.398
Gave different points of view	.818	−.107	−.207	−.158	.060	.375
Discussed current developments	.743	.035	−.142	−.270	.030	.340
Examination feedback valuable	.776	−.061	−.163	.336	.064	−.182
Examination methods fair	.808	−.149	−.163	.328	.069	−.146
Exams emphasized course content	.794	−.064	−.242	.289	.071	−.187
Readings/texts valuable	.639	.168	−.045	.131	.563	−.018
Added to course understanding	.754	.173	−.007	.122	.476	−.098
Course difficulty	.333	.840	.181	.075	−.068	.046
Course workload	.336	.793	.351	.037	.034	.065
Course pacing	.238	.794	.182	.092	−.100	.037
Hours/week outside class	.305	.726	.384	.068	.065	.115
Factor eigenvalues	20.67	3.95	1.75	1.42	1.18	1.05
% variance explained	59.1	11.3	5.0	4.0	3.4	1.9

Source: Abrami and d'Apollonia (1991), p. 414.

Figure 4. Nonrotated factor pattern matrix with six components extracted via principal components analysis for thirty-five items on SEEQ form. (from Abrami, P. C., d'Apollonia, S., and Rosenfield, S. (2007). The dimensionality of student ratings and instruction: What we know and what we do not. in R.P. Perry and J.C. Smart, (eds.), The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach. Springer, p. 421.

The SEEQ is a much longer form and includes several global measures. It's pattern with all items loading heavily into the first component is very similar to that yielded by CSUCI's Student Feedback on Course Learning Experiences form.

All indications on this preliminary study are favorable to the new ratings form being a viable tool. The final page of this report shows the report form which faculty volunteers received from the pilot initiative this spring, along with the raw data in Excel and all the original student written comments.

FACULTY XXX -Volunteer trial run

1. Organization & Planning	**3.64**
2. Clarity & Understanding	**3.58**
3. Perceived outcome or impact	**3.50**
4. Stimulation of interest in content	**3.62**
5. Encouragement and openness	**3.48**
6. Availability and helpfulness	**3.73**
Calculated Weighted Summary	**3.59**
Unweighted average all 20 items	**3.58**

Above you'll find six dimensions of teaching listed from top to bottom in order of general importance to learning. Also shown is a weighted overall summary calculated in accord with the relative importance of each dimension, and a simple average of scores on all twenty items. The scores on dimensions derive from averages of bundles of survey items shown on the graph below. The maximum possible score is a 4.0 in any category.

Also, check students' written comments for patterns that may validate particular dimensions. An easy way to look for patterns is to take students' positive comments and sort these according to six piles that follow the dimensions above. Then do the same with students' comments suggesting improvements. Patterns and scores on dimensions usually indicate good areas to target for improvement.

The occasional rude, inappropriate comment should be forgotten as quickly as possible, unless many define a dominant pattern. Should you choose to work on any dimension, see your faculty developer for help and resources.

ITEM DETAIL