Drawing conclusions: Using student drawings to inform teaching and learning in a university setting

Larry H. Ludlow
Damian Bebell

Boston College
Lynch School of Education
Department of Educational Research, Measurement and Evaluation

(Paper presented at the annual meeting of the American Educational Research Association, April 2003, Chicago, IL)

April 23, 2003

(Partially supported by a Boston College 2002 Faculty Teaching, Advising, and Mentoring Grant. Special thanks are extended to Ms. Katie Trong for her assistance and advice.)

Background

Understanding what makes a great teacher a great teacher is a long debated problem for those working in educational evaluation, policy, and measurement. The literature shows that from the very beginnings of public schools there have been measures in place to assess the quality of the teacher (Haney, Madaus, and Kreitzer, 1987). The earliest efforts were largely subjective observations and interviews carried out by untrained individuals. Surprisingly, many of these methods still thrive around the world (Haney, et al. 1987).

Since the release of *A nation at risk: The imperative for educational reform* in 1983 (National Commission on Excellence in Education, 1983), numerous observers have proposed reforms for education in the United States. In the intervening years, dozens of additional reports have called for reforms of areas of education ranging from teacher preparation and professional development to educational assessment and school governance.

Although it is beyond the scope of this proposal to reprise these reports, a common reform theme is that teachers within schools must become reflective practitioners if they are to become more successful in meeting the needs of increasingly diverse student populations. As Sternberg & Horvath (1995) wrote, “The current popularity of ‘reflective practice’ as a touchstone for teacher excellence suggests that, in the minds of many, the disposition toward reflection is central to expert teaching” (p. 15). It is thus not surprising that there is currently great interest in promoting reflective practice among teachers -- even though observers of the working lives of teachers, ranging from Philip Jackson (1968) to Michael Lipsky (1980), have noted that the beehive of activity in classrooms wherein teachers serve as supply sergeants, timekeepers, judges and in Lipsky's phrase "street level bureaucrats" leaves little time for reflection.

These continuous calls for reform have not been focused solely upon the traditional K-12 educational system. Recent reform efforts have been initiated for the teaching of college level statistics, particularly with respect to: cooperative learning (Garfield, 1993; Giraud, 1997); incorporating the ideas of context-relevant material (Sowey, 1995); devising alternative forms of testing and grading (Garfield, 1994); and team teaching the statistical software component of courses (Rumsey, 1998). However, documenting and gauging the extent to which such pedagogical changes have been effective in classroom practice is a major hurdle that remains to be addressed. Traditionally, end-of-course summative evaluation forms (required at most universities) are the "end-all, be-all" of measuring student experiences and teaching pedagogy of a course.

Without question these course evaluations have come to play a critical role in university decision making, but how useful a tool are they at helping professors better understand the perceptions of their students? The typical Likert-type evaluation form yields little information about the *student's experiences*—experiences that went beyond what was taught, how it was taught, and how well it was taught. Specifically, how did the student see, literally, the instructor, material on the board, fellow students, their own progress over the term, the level of interaction between classmates and the instructor, the salient teaching tools, the level of anxiety, boredom, and attention? How would the student describe the class to a friend?

The Imagine Scene

Imagine you have just presented a lecture on maximum likelihood estimation and the operation of the Newton-Raphson iteration technique to your doctoral students. You have presented this topic a half dozen times in your career. Tonight, however, you feel you have actually "taught" the students how

the technique works and why knowing about it is important. You even believe you have made the topic interesting, if not exactly exciting. Bottom-line, you have answered the ultimate questions, "So what?" and "Who cares?"

As you look at the students, awaiting their applause, a thought occurs, "What are *they* thinking?" You ponder this question as they file out of the room. There is no applause, no wave, no cheers, not even a "Nice job, Doc!" You wonder how they describe this class to their friends. What visual images do they construct for their audiences?

Over the next couple of days you ask a few students what they thought about the lecture.

- Did they understand it reasonably well?
- Was it clear?
- Did it make sense?
- Where were the tough parts?
- Where did they begin to lose it?

Their responses are non-descript--it was fine, it was interesting to see how the parts fit together, it made sense at the time, it was challenging but OK. Their responses, while somewhat supportive of your efforts, don't leave you satisfied. So you decide to try something unusual in the next class. At the start of the next class session you ask them to complete the following task:

> Think of a "typical" classroom teaching experience this semester with me. Now draw as best as you can, that classroom experience.

The class ripples with giggles. Students look at one another. Puzzled expressions are exchanged. Whispers and groans are heard. Some of them look at you as if you have gone really weird on them this time. Eventually they begin to draw.

Imagine your reaction to getting this drawing of "Yoda"

Show first drawing

This figure reflects the kind of positive classroom experience we wish we had on students every night we stand before them. The statements bursting from my head make sense. I come across as a person, even a personality-- "Yoda." There is a discernible message on the chalk board. The student doesn't even draw any stereotypical equations, numbers, curves or tables. Most dramatic of all, however, is the student's role. "Luke" is experiencing the "ah-ha" rush of insight that I have never seen represented as effectively as it is here.

Now, imagine your reaction to getting this drawing of "Scream"

Show next drawing

This figure is powerful and disturbing, even embarrassing. No words need to be written on this figure. No words written on a standard course evaluation form could adequately express the emotions of this student. This image is not one that we would like students to take away from our classes.

How do these drawings serve as a form of course evaluation?

My particular interest in teaching evaluations came about because of my dissatisfaction with a former dean's comments about the low ratings I received in a specialty course one spring, 10 years ago (Psychometrics). So I wanted to know, in general for all my courses, how my ratings looked over time—did they look like they were improving, dropping, or staying steady—regardless of what anyone else's looked like.

At BC, and probably at your institution too, student ratings are an integral part of the evaluation of teaching effectiveness and the results are used as one criterion for salary, promotion, and tenure decisions. One of the higher-level aims of this process is to improvement teaching but a systematic analysis of the evaluations can also serve a useful function when used to look at long-term career patterns.

Such an analysis might take into consideration:

- Personal changes in health and family circumstances, and
- Professional changes in tenure status, administrative demands, and teaching and research interests.

Unfortunately, these variables usually play little, if any role, in the analysis and interpretation of course evaluations. My interests, then, focused on how to make the evaluations more useful and valuable by devising a systematic approach to their analysis that could be shared with other faculty (Ludlow, 1996; Ludlow & Alvarez, 2001; Ludlow, 2003).

Kinds of standard quantitative analyses

We receive standard course evaluation summary sheets with the percent of students agreeing or disagreeing with a variety of classroom characteristics. These are sent to us on paper with no personal electronic records. There is no University maintained and useable data file.

To do any kind of analysis first requires that these percents, and other variables such as enrollment, be hand-entered into a data file. At present, the data file has expanded to include not only other results taken directly off the summary sheets but

- professional information such as tenure status (pre-post) and administrative status (chair-not chair) at the time the course was taught ,and

- personal information such as marital status (M-D-RM).

The data file contains the summary records of nearly 3000 student ratings of 95 classes taught over 19 years.

This information provides an extraordinary opportunity to analyze your own evaluations. In other words, what would you like to know about your overall course evaluation record? What about

- your ratings over time across the entire spectrum of what you've taught?
- a comparison of your undergraduate vs. graduate ratings?
- a comparison of your required vs. specialty courses?
- ratings pre to post tenure?
- ratings as a function of class size?

So, this is what I've doing to try to understand my student's ratings in order to make, if possible, teaching adjustments. Some of you may recall seeing some of these analyses over the past few AERA meetings.

(Show graphs)

Can we do something else based on these graphs? Yes, it is possible to build regression models that are now quite accurate at predicting evaluation ratings for current or upcoming classes. Now, as soon as I receive the latest summary of evaluation results I enter them into the data file and start looking at how those results add to the overall picture. The first question I ask is: "Have the new ratings strengthened or weakened previous patterns?"

Some examples of how I've used these results to improve teaching practice include:

- Continuing the use of small-group discussions and projects and the use of "what the hell did he say" nearest-neighbor moments.
- Continuous pressure to limit classes to 20 or less.
- Focusing on the presentation of principles and concepts along with factual information.
- Increasing my availability outside of class through stopping students in the hallways for discussion and encouraging email questions and answers.
- And, trying to keep a harmonious marital relationship.

How has this longitudinal statistical approach been of value to other faculty and administrators?

- It has helped some faculty build their own data bases for their own personal interests.
- It's helped others prepare their teaching portfolio for promotion and tenure, and annual review considerations, and
- this body of work has led to LSOE and University discussions of systematic use and review of evaluations.

The validity of this statistical modeling approach will be fully tested when longitudinal data covering 20 years of teaching by two other faculty members is finished being recorded and then analyzed.

What else can be done to understand the classroom experience?

Slowly a dissatisfaction grew with just the analysis of the quantitative summaries. They do not provide a rich enough source of information about what students experience in classes. These graphs are very effective for showing long term trends in ratings over one's career—also for comparing different types of courses and levels of students, and various other creative professional and personal characteristics. Ultimately, however, those graphs are all dependent upon one very simple piece of data—a filled-in circle on a standardized form. The question, then, is "what was driving the student to

make the mark they did—what were they re-experiencing about the course, what images came to mind to guide their mark, what lies behind the 1000's of numbers that produced those graphs?

My broader goal was to somehow get more personal information about what was going on in class. I believed that this type of information would yield insights in how to construct learning opportunities to enrich their statistics experience. Statistics is a tough topic for many students and anything that can address not only the professional presentation of it but their personal reaction to it, seems like a good thing to do.

In 1995, it came to my attention that elementary and middle school students were drawing interesting pictures of their classrooms that were useful to teachers for reflective purposes as well as to researchers for better understanding the student's experiences (Weber & Mitchell, 1995; Haney, et al, 1998). I decided to try the drawing technique in my classes as a means of obtaining additional course evaluation feedback.

In the early stages of the my project the students were simply asked to

> "Think of a "typical" classroom teaching experience this semester with me. Now draw as best as you can, that classroom experience"—this is the direction that produced the earlier drawings

At present, the instructions include:

(1) What visual image of a classroom experience comes to mind when you think of this course? Now draw as best as you can, that classroom experience. Include me, yourself, and anything else that represents for you that classroom experience. Ideally, someone else could look at your drawing and could then form a reasonable impression of your experience.

(2) On the back of your drawing write a full description of the scene you have drawn. Be as explicit, open, and comprehensive as you can.

(3) Finally, what "course evaluation" information does your drawing provide that your responses to the traditional scannable form do not contain?

Please try to accept my assurance to you that this information is confidential--I will not try to somehow figure out who passed in which one of these sheets. This information is part of a long-term research project that I am conducting on alternative modes of faculty evaluation assessment techniques.

I now have drawing data from one undergraduate course in research methods and seven different graduate courses in measurement, evaluation, and statistical analysis--Interpreting & Evaluating Research, Statistics I, Statistics II, Multivariate I, Multivariate II, Psychometrics, and Seminar in Educational Research. The courses cover seven years and all have been taught (and drawn) more than once. Some courses follow a traditional lecture format, others include a cooperative learning component. Most of the students are in the school of education. The lower level courses are required. Students in the higher level courses specialize in measurement, evaluation, and statistical analysis. There are now over 600 student drawings.

The "drawing evaluation" is performed immediately following their completion of the standard evaluation form. They are told that the drawing evaluation is part of a long-term research project and I encourage them to take the exercise seriously. They are also asked to write an arbitrary four-digit code of their choice on their scannable evaluation forms and their drawing.

What do the drawings look like?

To my amazement and delight, the drawings are rich beyond any expectation I held for them. Each drawing was initially looked at holistically within a class set. It was quickly discovered that this

type of exercise offered numerous opportunities to understand classes from the perspective of the students. The drawings were detailed, insightful, honest, and reflective. They provided a fascinating glimpse of what success or failure felt like; how the instructor was supportive or threatening; and, they provided an opportunity to understand how students perceived their peers as engaged and excited or bored and stressed. In comparison, the "voice" of the student seemed nearly absent from the traditional university course evaluation the instructor had been using for nearly 20 years.

Before looking at the drawings we can ask from a broad perspective:

- what is important in these drawings?
- what are students trying to convey about a particular course and instructor?
- what is unique and different about the courses?
- which patterns are similar across courses? and
- how can these drawings be systematically analyzed?

The drawings are presented to you as examples of sets of common ideas and themes as I see them. My interpretation of the drawings at this point has been entirely exploratory in that I did not start looking at the drawings with any initial hypotheses about what I would find and what the images would mean. I simply went from one drawing to the next , reacting to each as I looked at them, and started placing them in common piles.

(Show drawings)

Again, at this point, these observations are mine. It remains to be seen if others would draw the same interpretations and conclusions. Furthermore, the wider educational utility and impact upon teaching for other faculty remains to be established.

Analysis of the drawings

The broader research problem, of course, is how to analyze, interpret and explain not only these but other faculty initiated drawings in a way that is not self-serving, idiosyncratic, or arbitrary. Although student drawings helped inform me about my classroom practices and my wider classroom environment, the drawings needed to be analyzed in a more objective light for the purpose of quantitative comparisons and generalizability to other faculty who might choose to adopt this course evaluation tool. To this end issues of reliability and validity in the interpretation of these drawings are presently being addressed.

Accordingly, a method developed by Haney et al (Wheelock, Bebell & Haney, 2000) was adapted to develop an emergent analytic coding system. Such a coding method indicates whether individual drawings exhibit particular features. For instance, is the instructor depicted alone or with students; is he or she verbally addressing the class or writing on the blackboard; were computers, books, or projectors shown in use? To develop the emergent coding checklist two coders independently reviewed a sample of forty drawings and recorded the various features present in the drawings. The coders then compared their findings and condensed the list of features into a draft coding sheet.

This draft coding sheet was used by the same two raters to code a second sample of forty drawings. As the two raters worked independently, features were coded either present or absent. In addition, the raters took notes of features that existed in the drawings but were absent from the coding sheet. The codings for these two raters were then compared. For features that had high levels of

agreement, formal descriptions (operational definitions) of each feature were developed. For features that had low levels of agreement, the coders worked together to examine drawings for which there were discrepancies to identify reasons for discrepancies and to develop an operational definition of the feature. If a common definition of the feature could not be developed, the feature was removed from the list and the coding system.

Each drawing was coded using the above list and was recorded directly into a Microsoft Excel 98 spreadsheet. Additionally, some limited information about the student artist (course, semester, year, level of study, etc.) was recorded. This allows for the reporting of how drawings depict systematic changes across courses and time. For example, it is possible to show how classroom interactions changed in a research methods course when small-group discussions and in-class projects were introduced. In addition, it is possible to address questions like: "Are students depicting themselves as more anxious and confused in the higher-level specialty statistics courses or in the lower-level required statistics courses?" This presentation will illustrate the efficacy of using drawings to answer such questions.

We have just completed the coding of all drawings. The categories and their operational definitions for coding presently consist of:

Instructor Presence/Affect:

Instructor is present: Instructor is somehow somewhere depicted in the drawing.
Depicted Positively: Instructor's facial expression is positive (smiling), or depiction of positive speech (praise, support).
Depicted Negatively: Instructor's facial expression is negative (frowning), or depiction of negative speech (confusion, malice).
Depicted Neutrally: Instructor face and speech (if any) are not positive or negative but visible.
Can't discern affect: Instructor is present but affect is not visible.

Instructional location of instructor

At board: Instructor is drawn located at or near the board (not necessary actively using it).
At overhead: Instructor is located at/or near overhead projector (not necessary actively using it).
With student(s): Instructor is drawn EITHER physically with student(s) or metaphorically with student(s)

Instructor Interaction
Instructor Speaking: Words/phrases are depicted coming from Instructor.
Supportive: Instructor is speaking (see above) words of support or encouragement.
Asking Question: Instructor is posing a question. (student's my be present or not present).
Speaking Statement: Instructor is lecturing/instructing
Instructing single student: Instructor is addressing one student. (words or phrases need not be present).
Instructing class/lecturing: Instructor is addressing(?) more than one student (words or phrases need not be present).

Student(s) Present: Figures are present in the drawing other than instructor or TA.
1 depicted: Only 1 student is depicted anywhere in the drawing.
2 or more: 2 or more students are depicted anywhere in the drawing.
Sitting in groups: More than 1 student is located in a cluster of desks or with other students.
Sitting in rows: Students are arranged in rows or columns (either in desks or not).
Asking a question: 1 or more student is actively speaking a question, thinking a question.
Answering a question: It is obvious that 1 or more students are actively speaking a response to a question or thinking a response to a question.
Student to student interaction: There is some evidence of communication and discussion Between (amongst) students (may include arrows connecting students or thoughts) (Not just sitting in groups).

Student Depiction
Depicted Positively: At least 1 student's facial expression is positive (smiling), or positive speech.
Enthusiasm/excited: 1 or more student is obvious depicted as being enthusiastic or excited.
Depicted Negatively: At least 1 student's facial expression is negative (frowning), or negative speech.
Confusion, frustration or stress: 1 or more student is obvious depicted as confused, frustrated or stressed.
Depicted Neutrally: Face and speech of students is visible but not positive or negative.
Can't discern affect: Students are present but affect is not visible.

Course Experience(s)
AHA/light bulb/lightning bolt: At least 1 student depicts a light bulb, or AHA!, or lightning bolt
Other sudden insight: Some other evidence of the sudden understanding of a concept or

idea
Understanding over time: Some depiction of before/after learning, gradual growth of learning or understanding
Enthusiasm/excited: The student(s) is obvious depicted as being enthusiastic or excited.
Sleeping/bored: Student(s) is depicted sleeping or obviously bored.
Daydreaming: Student(s) is obviously daydreaming (off topic thought bubbles, etc.)
Crying: 1 or more students are depicted as crying as evidenced by text or tears.
Angry: 1 or more student is expressing hostility, "pissed off"
Sleeping/bored: 1 or more student is depicted as sleeping or obviously bored

About classroom
Computer depicted: A computer is present somewhere in the drawing
Overhead projector depicted: An overhead projector is present in the drawing
Laser pointer: A laser pointer (or beam) is present in the drawing
Clock: Clock or representation of a clock is depicted somewhere in the drawing.

Other
Readable text: There is readable instructor related text somewhere in the drawing (words or sentences, NOT SPEECH—typically board work).
Unreadable text: There is unreadable instructor related text somewhere in the drawing (scribbles, etc. NOT SPEECH—typically board work).
Statistical symbols/formula: Stats or formula are located somewhere in the drawing.
Graphical representation of the data: Statistical graphs or figures are located somewhere in the drawing.
Assistants present: One of Instructor's graduate assistants are present (Chris, Julie, Jere, Camelia).
Students thought(s) depicted: Words, thoughts or ideas are presented in the drawing representing the student's thoughts.
Metaphorical: The drawing uses a metaphorical representation of the classroom or experience, rather than a pictorial depiction of the actual classroom environment.

The codes for about ½ of the drawings based on these characteristics have now been entered into a data file. The initial results from the comparisons of the codes are now available. Table 1 contains the percents for the various coding categories. Three different intact course offerings are included along with the summary across this sample of 292 drawings. The percents have been highlighted to focus attention on particularly interesting aspects of these initial results.

For example, to form a general impression of my courses I can look down the "Overall Drawings" column and say that the instructor tends to be drawn positively as he spends most of his time

at the board and his teaching approach is primarily lecturing to students sitting in traditional rows who are generally depicted positively although there is evidence that many students see themselves or others as being confused and overwhelmed at times.

As for differences in these classes themselves we also see some interesting tendencies. These three course were specifically chosen because we believed that they the experiences of students in these courses is very different. The undergraduate course is a research methods course that has no statistical component—it consists of frequent discussions of published articles. The Intermediate statistics course is required of all doctoral students and is generally perceived as being difficult. The Psychometrics course is a specialty course for the measurement, evaluation and statistical analysis program students.

One of the first things to notice is that the use of the chalk board tends to decrease as the technical nature of the course increases while, correspondingly, the use of the overhead projector for technical detail and computer output tend to increase. While the two lower level courses show some evidence of asking and answering questions posed by the instructor, there is no such evidence in these drawings of this happening. The representation of technical detail in the drawings increases with the level of the course as does the extent to which student thoughts are depicted.

Show results

Linkage of drawings to the ratings

At this time, the remainder of the drawing codes are being entered in the data file. It will then be possible to get summary statistics such as these for each separate class that has been taught in the phase of the research. These summary statistics will then be entered into the course evaluation data file

described earlier that has the quantitative ratings by the students. It will then be possible to link the quantitative course summaries with these qualitative summaries.

Furthermore, it will be possible to combine the qualitative characteristics of the individual student drawings with the individual student quantitative course evaluation ratings. This step will be possible because students created a unique four or five digit identification code that they wrote on their drawings and their original scannable evaluation sheets (the formal evaluation sheets are returned to us). This linkage will allow for the creation of a datafile that has both their drawing codes and their actual evaluation ratings. It will then be possible to link the summary course evaluation results to the individual student drawings and evaluations.

Form of Course Evaluation Effectiveness

Included with the drawing exercises were instructions to describe the scenes and to offer their insights into what their drawings have to offer in the way of course evaluation information that the scannable forms were not able to convey. The following fragments of statements by students illustrate the point of this part of the research:

- This type of evaluation gets to the heart of the class-what the atmosphere is really like and what was important for them in their learning.
- provides more information on teaching style and student interaction.
- I could be more descriptive and honest in this evaluation.
- the regular course evaluation does not ask if the Prof really cared about the kids learning. everything and to what degree.

Summary:

I've been trying to understand how students experience classes and how those experiences affect learning ever since I first started teaching. For example:

- What techniques of engagement or interaction work well
- What delivery systems (overhead, handouts, board work) work well
- What do students pay attention to during class
- What do they think about the overall class environment
- How do they experience the class and describe it to others.

Basically, I've questioned how different professional and personal characteristics affect one's teaching and the extent to which those variables affect their learning and the extent to which those variables and their affects are alterable. To answer these kinds of questions, I've been working on ways to extract more useful personal classroom experience information from students.

One solution has been to develop an intensive statistical investigation into the understanding and modeling of student response to standard course evaluations. The systematic, longitudinal analysis of evaluations presented here has been a useful means of offering insight into what factors underlie the ratings students provide. Such an analysis can be particularly powerful and contextually relevant when would instructor specific professional and personal variables are included.

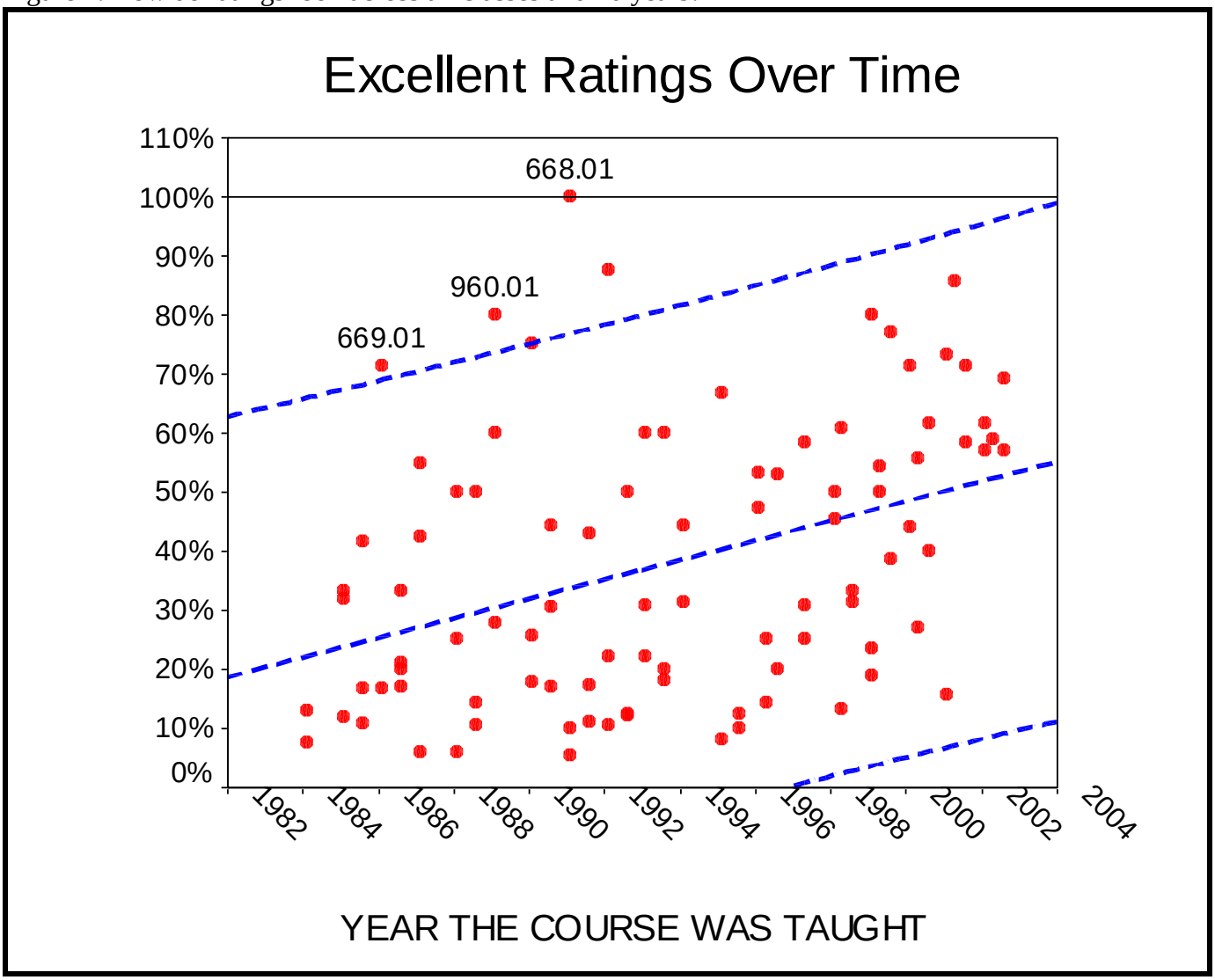
In addition to this relatively straightforward approach, the classroom drawings can also be useful at extracting more emotional, affective student experiences—both in terms of the structural aspects of instruction and how that instruction had an impact on the student. These drawings are extremely personal, emotional, and powerful. On one hand, they can bring us pleasure and satisfaction—and those are good feelings for teachers to experience—pay certainly isn't the only reason we do what we do. On the other hand, they can be profoundly disturbing and haunting. Which ever is the case, they will affect

you, and potentially your practice, in ways that will never be possible from looking at a summary rating on a computer printout.

As educational researchers, we are well aware of the negative attitude and belief systems that many students carry as baggage to required statistics courses. It is clear that these attitudes and beliefs are not necessarily well communicated on standard course evaluation forms. Even more of a problem from the instructor's standpoint is the inability of students to effectively articulate changes in negative attitudes and beliefs.

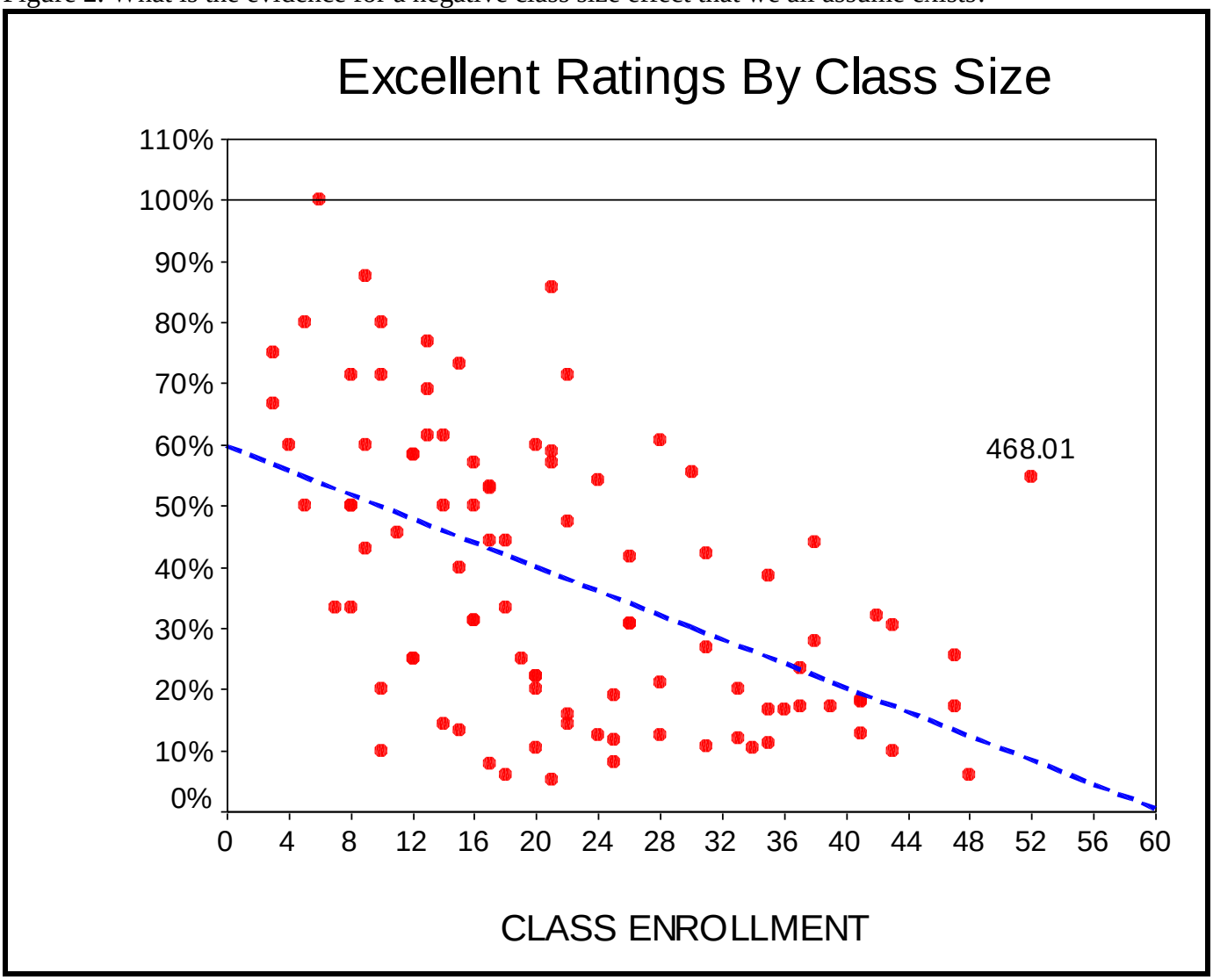
The systematic analysis of these drawings, combined with the opportunity to link them to the standardized ratings from the same students, has created a unique approach to the assessment, interpretation, and evaluation of instructor and course effectiveness. Drawings depicting scenes with "ah-ha's" (see Figure 1), dead-fish expressions, confusion (See Figure 2), light bulbs turning on, gibberish on the blackboard, celebrations on a mountain top, and tear-drops on an anguished face effectively communicate what students feel in classes. Those expressions lead to an intensity of self-reflection about practice that is impossible to ignore and is virtually impossible to experience with any standardized course evaluation now in use.

Figure 1. How do ratings look across all classes and 20 years?



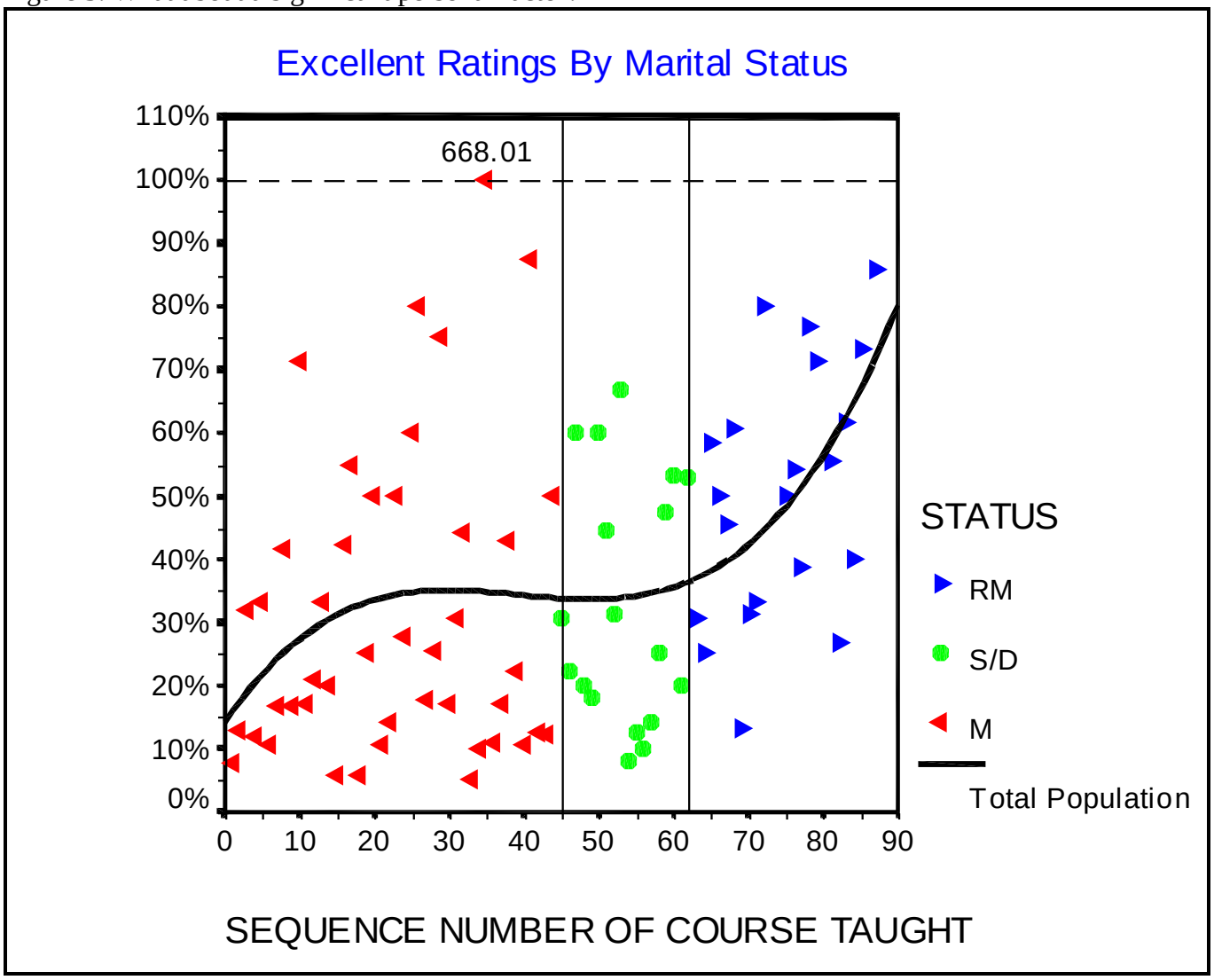
The ratings follow a general upward trend over time across all classes. The center dashed line is the regression line—the predicted excellence rating for a given year the course was taught. The other two dashed lines represent the 95% confidence interval around the regression—the region within which we most of the ratings to lie. Note that at the time that 668, 669, and 960 were taught for the first time (.01), their ratings were much higher than expected based on other classes up to that point.

Figure 2. What is the evidence for a negative class size effect that we all assume exists?



There is a clear, unmistakable relationship between the ratings and class size—as the enrollment increases, the ratings tend to drop. The drop is actually at the rate of a decrease of 1% in excellence rating for each additional student added to the class. Note that 468.01 is an outlier—it was the first time it was taught, it attracted a crowd, and it differently substantially from the way it had been previously taught. This is what we take to the Dean's office to argue for smaller classes!

Figure 3. What about a significant personal factor?



The term "spillover-effect" is usually applied to the spillover of pressures from work to the home. Here, it is used to refer to spillover from home to work. Specifically, during the early phase of marriage and work at BC the ratings show an upward trend. The ratings, however, start to fall off prior to and continuing into the period of separation and divorce. During this period they again change direction and begin to recover prior to and continuing into the current remarried stage.

Table 1: Percent of drawings representing various classroom experiences

	Research Methods (UG)	Intermediate Statistics	Psychometrics	Overall Drawings
Instructor Presence/Affect:				
Depicted Positively	0.57	0.30	0.36	**0.42**
Depicted Negatively	0.00	0.10	0.00	0.02
Depicted Neutrally	0.14	0.25	0.45	0.26
Location of Instructor:	0.00	0.00	0.00	
At board	**0.77**	**0.70**	**0.64**	**0.65**
At overhead	**0.26**	**0.35**	**0.36**	0.17
With student	0.11	0.20	0.00	0.24
Instructor Interaction:				
Instructor Speaking	0.43	0.35	0.45	0.37
Supportive	0.03	0.05	0.00	0.10
Instructing class/lecturing	0.83	0.80	0.82	**0.68**
Student(s) Present:				
1 depicted	0.20	0.20	0.00	0.23
2 or more	0.80	0.75	0.91	**0.71**
Sitting in groups	0.03	0.15	0.00	0.16
Sitting in rows	0.80	0.60	0.82	**0.57**
Asking a question	0.23	0.30	**0.00**	0.14
Answering a question	0.03	0.00	**0.00**	0.02
Student to student interaction	0.09	0.25	0.00	0.13
Student Depiction:				
Depicted Positively	0.23	0.20	0.00	**0.27**
Depicted Negatively	0.06	0.15	0.09	0.15
Depicted Neutrally	0.00	0.20	0.18	0.12
Course Experience(s):				
AHA/light bulb/lightning bolt	0.06	0.00	0.00	0.08
Other sudden insight	0.06	0.05	0.00	0.05
Understanding over time	0.03	0.05	0.00	0.11
Enthusiasm/excited	0.03	0.00	0.18	0.11
simple understanding	0.06	**0.50**	0.00	0.23
daydreaming	0.09	0.00	0.00	0.04
confused/overwhelmed/lost	0.17	0.50	0.55	**0.33**
About Classroom:				
Computer depicted	0.03	0.00	0.09	0.07
Overhead projector	**0.29**	**0.40**	**0.45**	0.22
Laser pointer	**0.26**	**0.30**	**0.45**	0.14
Clock	0.03	0.00	**0.18**	0.04
Other:				
Readable text	0.34	0.40	0.36	0.24
Unreadable text	0.31	0.45	0.09	0.25
Statistical symbols/formula/tables	**0.11**	**0.45**	**0.73**	0.27
Graphical representation of the data	0.23	0.30	0.18	0.27
Students thought(s) depicted	**0.26**	**0.30**	**0.36**	0.32
Metaphorical	0.00	0.10	0.00	0.16
n=	35	20	11	292

References

Garfield, J. (1993), Teaching Statistics Using Small-group Cooperative Learning, *Journal of Statistics Education* [Online], 1(1).

Garfield, J.B. (1994), Beyond Testing and Grading: Using Assessment to Improve Student Learning, *Journal of Statistics Education* [Online], 2(1).

Giraud, G. (1997), Cooperative Learning and Statistics Instruction, *Journal of Statistics Education* [Online], 5(3).

Haney, W., Russell, M., Gulek, C. & Fierros, E. (Jan/Feb. 1998). Drawings on education: Using student drawings to promote middle school improvement. *Schools in the Middle: Theory into Practice*, 38-43.

Haney, W., Madaus, G. and Kreitzer, A. (1987), Charms talismanic: Testing teachers for the improvement of American education. In E. Rothkopf (Ed.) *Review of Research in Education*. Volume 14, pp. 169-238.

Jackson, P. (1968). *Life in classrooms*. New York: Holt, Rinehart & Winston.

Lipsky, M. (1980). *Street-level bureaucracy : Dilemmas of the individual in public services*. New York : Russell Sage Foundation.

Ludlow, L. H. (2003). Rethinking practice: Using faculty evaluations to teach statistics. *Journal of Statistics Education*10(3). www.amstat.org/publications/jse/v10n3/ludlow.html.

Ludlow,L.H. (2002). A structural model for understanding faculty evaluations. Paper presented at the American Educational Research Association annual meeting. New Orleans, LA. April 5.

Ludlow,L.H. & Alvarez-Salvat,R. (2001). Spillover in the academy: Marriage stability and faculty evaluations. *Journal of Personnel Evaluation in Education, 15:2*, 111-119.

Ludlow, L.H. (1996). Instructor evaluation ratings: A longitudinal analysis. *Journal of Personnel Evaluation in Education, 10*, 83-92.

Rumsey, D.J. (1998), A Cooperative Teaching Approach to Introductory Statistics, *Journal of Statistics Education* [Online], 6(1).

Sowey, E.R. (1995), Teaching Statistics: Making it Memorable, *Journal of Statistics Education* [Online], 3(2).

Sternberg, R. J & Horvath, J. A. (1995). A prototype view of expert teaching. *Educational Researcher* (24:6), pp. 9-17.

Weber, S. & Mitchell, C. (1995), *That's Funny, You Don't Look Like a Teacher*, Washington, D.C.: Falmer.

Wheelock, A., Bebell, D., & Haney, W. (2000). "What Can Student Drawings Tell Us About High-Stakes Testing in Massachusetts?" *TCRecord Online* www.tcrecord.org