UNIversitas: Journal of Research, Scholarship, and Creative Activity

Volume 5
Number 1 *Forum Theme: Truth, Faith, and Intelligent Design*

Article 4

3-2009

Using the DELPHI Method to Collect Feedback on Student's Perceptions of Teaching Quality

Otto H. MacLin
University of Northern Iowa

M. Kimberly MacLin
University of Northern Iowa

M. Catherine DeSoto
University of Northern Iowa

Robert T. Hitlan
University of Northern Iowa

John E. Williams
University of Northern Iowa

Volume 5 | Issue 1 ISSN | 1558-8769

UNIVERSITAS

Using the DELPHI Method to Collect Feedback on Students' Perceptions of Teaching Quality

Otto H. MacLin, M. Kimberly MacLin, M. Catherine DeSoto,
Robert T. Hitlan, & John E. Williams

In educational institutions, teaching effectiveness is a highly valued asset among administrators, professors, and students alike. Information gathered from students is often used as a basis for promotion and tenure decisions (Abrami & d'Apollonia, 1999; Waters, Kemp, & Pucci, 1988), and, ideally, formative purposes. However, students do not always believe that their evaluations carry much weight (Chen & Hoshower, 2003; Spencer & Schmelkin, 2002). This is likely due to the fact that summarized results from student evaluations often do not get in the hands of professors until after that particular course has concluded and, therefore, the feedback does not directly benefit the students who provided it. The goal of the current paper is to introduce a method that can be used by instructors to evaluate the effectiveness of their teaching in a particular course in such a way as to implement change in the course if necessary for those very same students. First, we will discuss teaching effectiveness in general; second we will introduce the DELPHI method and its usefulness in evaluating effective teaching; and third we will report on the results of using this method in our courses with the goal of improving the learning experience for the students providing the feedback.

An important first step to being an effective teacher is being familiar with the extensive literature base available on effective teaching. Even defining effective teaching is not an easy task. The simplest definition (while somewhat cynical) is that effective teaching is anything that results in positive evaluations of teaching (Neath, 1996; Nussbaum, 1992). Many researchers have conducted studies to uncover what qualities and corresponding behaviors make for effective teaching e.g., Buskist, Sikorski, Buckley, & Saville, 2002; Epting, Zinn, Buskist, & Buskist, 2004). Effective teaching is complex and research indicates that measures of effective teaching are multifaceted and multidimensional (Marsh & Roche, 1997; Sheehan & DuPrey, 1999; Tang, 1997).

Previous research has found that effectiveness is related to physical attractiveness and vocal clarity (Feeley, 2002), teacher likeability and interpersonal interactions, a positive experience (Delucchi & Pelowski, 2000; Sinai, Tiberius, de Groot, Brunet, & Voore, 2001), teaching style (McKeachie, Lin, Moffett, & Daugherty, 1978), teacher extroversion and age (Radmacher & Martin, 2001), humor (Kher, Molstad, & Donahue, 1999), proper workload (Marsh, 2001), clear presentation of the material and preparedness of the instructor (Carkenord & Stephens, 1994; Tang, 1997), rapport (Lowman & Mathie, 1993; Perkins,

Schenk, Stephan, & Vrungos, 1995), and encouragement of questions (Carkenord & Stephens, 1994).

Schaeffer, et al., (2003) found that of the factors related to teaching effectiveness approachability, creativeness and interest, encouragement and caring, enthusiasm, flexibility and open mindedness, knowledge, realistic expectations and fairness, and respectfulness ranked at the top. Feldman (1976) identified teacher's interest, knowledge, public speaking skills, value of the course material, and intellectual expansiveness as important elements to effective teaching. Jackson et al. (1999) found that rapport with students, course value, course organization, fairness in grading, difficulty of the course, and course workload for the students were key indicators of teaching effectiveness.

Although it may be difficult to define effective teaching, it is a construct that is stable, with a high degree of agreement among students (Harrison, Ryan, & Moore, 1996) and instructors (Miller, Dzindolet, Wienstein, Xie, & Stones, 2001; Schaeffer, Epting, Zinn, & Buskist, 2003).The goal of this paper is not to detail every factor that contributes to effective teaching (there are many), but rather, to propose a method for evaluating what works and what does not in the teaching environment you are creating.

The DELPHI Method

The DELPHI method was developed by the RAND Corporation in the late 1950s, and uses an organized procedure of polling experts on a topic of interest (Gordon & Helmer, 1964; Helmer & Rescher, 1958). Researchers have used the DELPHI survey to examine how supervisors make treatment decisions (Kessler, Nelson, Jurich, & White, 2004), assessment of occupational and family therapy practices (Deane, Ellis Hill, Dekker, Davies, & Clarke, 2003; Jenkins, 1996; Jenkins & Smith, 1994), perceptions of quality of life (Meuleners, Binns, Lee, & Lower, 2002), perceived risk (Moldrup, Morgall, & Almarsdottir, 2002) and the development of questionnaires (Gaskin, O'Brien, & Hardy, 2003; Spangenberg & Theron, 2002). Its procedures are ideally suited for studying teaching effectiveness as well.

The basic DELPHI method is a two- and sometimes three-round process (Linstone & Turoff, 2002). In the first round, researchers identify participants based on their expertise and their potential contribution. For example, if we wanted to know what makes a toy fun, we might use children as our experts. We might ask a very general question such as "tell me 10 things that make a toy fun." We could also ask what makes toys not fun. This process is divergent because we expect to generate a variety of responses generated in isolation of the other experts, thus avoiding a groupthink mentality that enables participants to express their opinions freely (James, Aitken, & Burns, 2001). Round 2 is convergent. Responses from Round 1 are compiled and grouped into like sets. For example, the response "a fun toy doesn't break" could be combined with "they are hard to break" into a category of "Unbreakable" as an attribute of a fun toy. Categories are then compiled into a single list of

all items. This list is then presented back to the same group of experts for controlled feedback. The experts are asked to indicate "what makes a toy fun" by checking off as many items that apply. The benefit of Round 2 is that participants have a second opportunity to respond as they did in Round 1 or they can modify responses if something appears on the list that they had not considered previously. Responses from this list are now rank ordered to determine what are the most important attributes of a fun toy. During an optional Round 3, a survey can be developed based on the responses generated in Round 2. This survey then can be re-administered to the experts, or if desired, to a new group of participants.

Our goal was to use the DELPHI method as a barometer for evaluating perceived teaching quality within a particular class during the course of a semester. This method not only helps uncover what constitutes effective teaching, but more to the point of this paper, evaluates specific teaching behaviors within a specific teaching context. This information can then be used to make changes within that context to the immediate benefit of that particular group of students.

Method

Participants

An upper-level class of 65 psychology students at the University of Northern Iowa participated for partial course credit. Sixty three percent (63%) of the students were female.

Procedure

Round 1. We asked participants to think of the most *effective* teacher they have had in the past and to write down what made that teacher effective. We then asked them to think of the most *ineffective* teacher they had ever had and to write down what made him or her ineffective. We tabulated responses into a list format for use in the second round. In all, the experts generated 302 responses for describing a highly effective teacher and 246 responses when describing a highly ineffective teacher.

Round 2. We categorized responses from Round 1 into a new list consisting of two general categories: effective and ineffective teaching. These categories contained 28 separate items for a highly effective teacher and 19 items for a highly ineffective teacher (see Table 1). We asked the students to select between 5 and 10 items from each list.

Table 1: Round 2 Frequency Counts For Highly Effective and Highly Ineffective Teachers

Highly Effective Teacher:		Highly Ineffective Teacher:	
44	Approachable	51	Unapproachable
38	Humorous/Fun	44	Boring
34	Enjoys Material/Excited About Teaching	39	Intimidating/Jerk
32	Gives Good Examples	35	Lectures Not Related To Test
31	Good Communicator	35	No Sense Of Humor
		32	Didn't Explain

31	Organized/Prepared
28	Knowledgeable
28	Remembers What It Is Like To Be A Student
26	Fair/Reasonable Standards
23	Lectures are relevant
22	Flexible
21	Reviews
20	Willing To Help
18	Answers Questions
17	Positive
16	Respectful
14	Energetic
13	Materials Interesting
12	Understanding
11	Easy Going
11	Teaches With Variety
9	Available
9	Patient
3	Gives Group Work
2	Caring
1	Know Their Students
1	Gives Clear, Concise
1	Definitions Of Topic/Terms
32	Bad Communication/ Too Fast – Too Slow
30	Monotone Voice
26	Can't Answer Questions
26	Unprepared
24	No Energy
21	Strays From Subject
18	Strict
5	Lectures From Text
1	Makes Us Take All These Notes, Then Says It Won't Be On The Test
1	Assigns Group Work/ Projects
1	Demeaning To Students/ Cuts Students Down
1	Unfair Testing Strategies (Correct Answers Are Her/His Opinion)
1	Doesn't Care About Students' Lives/Feelings

Results and Discussion

Results indicated that a highly effective teacher is an *approachable, humorous, fun person who is excited about teaching*. The highly ineffective teacher is an *unapproachable* and *boring jerk*, and in general, is opposite the effective teacher (see Table 1). Also, items such as *being caring, knowing the student*, or *teaching from the text* and *making the students take notes that will not be on the test* are not as important to the majority of the students, although still important to some. Even though these items are about teachers in general, they can be used immediately as a barometer to ensure that teachers are not exhibiting behaviors that are undesirable to current students.

While some global traits may be useful to change across all courses (e.g., *talks too fast*) others, might be class-specific referring to a specific interaction or event in the course (e.g., *being disrespectful*). This method can help tailor teaching style to the particular audience.

In sum, having knowledge of what traits comprise effective teaching in the eyes of students is important and relatively easy to obtain using the DELPHI method. The DELPHI method is anonymous, easy to administer and allows for distilling abstractions such as effective teaching into something manageable and directly applicable.

Additionally, one could use the Round 2 list to develop an instrument and administer that to students for further information about the perceptions of the teaching in particular. This is what we did in Study 2.

Study Two

Participants

Undergraduate psychology students (N = 320) from 6 courses participated at the University of Northern Iowa. Course size ranged from 17 to 181. These measures were administered in similar fashion as traditional teaching evaluations, thus, demographic information is not available. The demographic breakdown of psychology courses like these courses is typically 60% female, 90% White, with an age range of 18-22.

Materials and Procedure

A 30-item survey was developed using the DELPHI method. Study one's list was used (see Table 2) with the following additions: GPA, and a 'happiness' question (asking them to rate on a scale of 1 to 7 where 1 = "*I am happy with the class*" and 7 = "*I am unhappy with the class*"). We administered the surveys around the middle of the semester.

Table 2: Likert Survey used in Round 3 of DELPHI Study

Approachable	1	2	3	4	5	6	7	Unapproachable
Humorous/Fun	1	2	3	4	5	6	7	No Sense of Humor
Enjoys Material	1	2	3	4	5	6	7	Hates Material
Likes Teaching	1	2	3	4	5	6	7	Hates Teaching
Good Examples	1	2	3	4	5	6	7	Poor Examples
Good Communicator	1	2	3	4	5	6	7	Poor Communicator
Organized/Prepared	1	2	3	4	5	6	7	Unorganized/Unprepared
Knowledgeable	1	2	3	4	5	6	7	Unknowledgeable
Relates to Students	1	2	3	4	5	6	7	Can't Relate to Students
Sets Fair Standards	1	2	3	4	5	6	7	Sets Unfair Standards
Lectures are Relevant	1	2	3	4	5	6	7	Lectures are Irrelevant
Flexible	1	2	3	4	5	6	7	Inflexible
Reviews	1	2	3	4	5	6	7	Doesn't Review
Willing to Help	1	2	3	4	5	6	7	Unwilling to Help
Answers Questions	1	2	3	4	5	6	7	Can't Answer Questions
Positive	1	2	3	4	5	6	7	Negative
Respectful	1	2	3	4	5	6	7	Disrespectful
Energetic	1	2	3	4	5	6	7	Not Energetic
Materials interesting	1	2	3	4	5	6	7	Materials Uninteresting
Understanding	1	2	3	4	5	6	7	Not Understanding
Easy Going	1	2	3	4	5	6	7	Strict
Varies Teaching	1	2	3	4	5	6	7	Doesn't Vary Teaching
Available	1	2	3	4	5	6	7	Unavailable
Exciting	1	2	3	4	5	6	7	Boring
Not Intimidating	1	2	3	4	5	6	7	Intimidating

Explains	1	2	3	4	5	6	7	Doesn't Explain
Talks just Right	1	2	3	4	5	6	7	Talks too Fast
Talks just Right	1	2	3	4	5	6	7	Talks too Slow
Good Voice	1	2	3	4	5	6	7	Monotone Voice
Keeps on Subject	1	2	3	4	5	6	7	Strays from Subject
Effective	1	2	3	4	5	6	7	Ineffective
Happy	1	2	3	4	5	6	7	Unhappy

Results and Discussion

The primary use of Round 3 of the DELPHI method was to administer the survey that was developed with the experts in Rounds 1 and 2, to similar experts to gain feedback upon actual performance (as opposed to general impressions of good and bad performance). Therefore, we first will report results from one course to illustrate how the instructor could use that information, mid semester to continue with what works, and modify what does not. Then, we will report some findings across all courses surveyed.

Evaluating the data from one course (Introductory Psychology, N=181) provides information for that instructor on specifically what is working and what is not (See Table 3). A ranking means analysis is presented. Low scores represent higher quality on that factor (refer to Table 2 for the instrument and scaling).

Table 3: Survey results from Instructor A

Item	Mean	SD
Reviews	2.64	1.41
Varies Teaching	2.59	1.38
Talks Right	2.55	1.38
Exciting	2.10	1.16
Flexible	2.09	1.17
Not Intimidating	2.06	1.26
Available	2.05	1.09
Talks Right Speed	1.99	1.05
Organized / Prepared	1.92	1.07
Fair	1.91	1.08
Approachable	1.91	1.17
Material is Interesting	1.91	1.08
Willing to Help	1.87	1.09
Understanding	1.85	1.03
Tests Relate	1.81	1.04
Easy Going	1.80	0.96
Happy	1.78	1.16
Humorous	1.73	0.92
Explains	1.72	0.93
Good Communication	1.70	0.96
Keeps on Subject	1.67	0.80
Effective	1.64	0.88
Voice Just Right	1.59	0.75
Lectures are Relevant	1.57	0.88
Uses Good Examples	1.57	0.82
Answers Questions	1.56	0.89
Respectful	1.50	0.78
Enjoys Material	1.50	0.84
Likes Teaching	1.48	0.83
Energetic	1.46	0.76
Positive	1.45	0.78
Knowledgeable	1.40	0.69

Upon review, we can see that no means are above the midpoint of the scale, indicating overall good quality teaching. However, certainly, some areas are better than others, with *knowledgeable*, *positive* and *energetic* receiving the best marks, and *reviews for the exams*, *varies teaching style* and *talks just right* receiving the poorest marks. This provides detailed, quantitative feedback to this instructor on what could be improved in that particular course and changes could be made mid-semester. These changes make the course better for the students who had the concerns, as opposed to having to wait to implement them for a future class because you did not receive the feedback until after the course was over.

Compare the previous situation with another instructor and class (N=19) where there is more need for improvement (See Table 4). Here we can see that *keeps on subject* and *knowledgeable* are rated favorably, but that several items are above the mid-point of the scale and could use improvement. Thus, giving this DELPHI prior to formal departmental or university evaluations, can give the instructor valuable feedback that allows for changes during the semester to benefit the students taking the course who provided the feedback, but also aids the instructor in potentially improving his or her teaching style prior to being formally evaluated.

Additionally, we were interested in evaluating the data across classes. We found that 'happiness' with the course (recall that students were asked, "*Overall, are you happy with the class?*") was significantly related to effectiveness as measured with effectiveness as measured by Q30 of our DELPHI survey $r\,(310) = .80, p < .001$.

Further, do students with higher GPAs evaluate teaching effectiveness and effectiveness factors differently? The students in this study had a mean GPA of 3.22 and a median of 3.30. GPA was significantly correlated with teaching effectiveness $r = .138, p < .05$, as well as 22 of the 31 other items in our DELPHI survey.

Table 4: Survey Results from Instructor B

Item	Mean	SD
Varies Teaching	3.68	1.80
Humorous	3.63	1.83
Exciting	3.58	1.57
Approachable	3.53	1.87
Easy Going	3.42	1.77
Flexible	3.32	1.77
Tests Relate	3.32	2.00
Material is Interesting	3.26	1.69
Good Communication	3.16	1.74
Willing to Help	3.16	2.09
Not Intimidating	3.11	1.91
Available	3.00	1.70
Understanding	3.00	1.67
Likes Teaching	2.89	1.79
Happy	2.84	1.83
Answers Questions	2.79	1.78
Enjoys Material	2.53	1.39
Positive	2.47	1.50
Respectful	2.37	1.38
Energetic	2.37	1.50
Fair	2.32	1.53
Reviews	2.32	1.80
Voice Just Right	2.16	1.12
Effective	2.16	1.34
Explains	2.05	1.27
Uses Good Examples	2.00	1.25
Talks Right	1.68	0.67
Talks Right Speed	1.68	0.67
Lectures are Relevant	1.63	0.96
Organized / Prepared	1.63	1.12
Knowledgeable	1.37	0.60
Keeps on Subject	1.26	0.45

In conclusion, the DELPHI method can be an effective means of evaluating the instructor's performance during the semester. Because the DELPHI method helps to identify specific factors of interest, it allows the instructor to target specific behavioral factors relevant to the course that are in need of improvement. This in turn provides a better learning environment, potentially improving teaching effectiveness and teaching effectiveness ratings on formal class evaluations.

References

Abrami, P. C., & d'Apollonia, S. (1999). Current concerns are past concerns. *American Psychologist, 54*, 519-520.

Buskist, W., Sikorski, J., Buckley, T., & Saville, B. K. (2002). Elements of master teaching. In S. F. Davis & W. Buskit (Eds.). *The teaching of psychology, Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 27-39). Mahwah, NJ. Lawrence Erlbaum Associates, Inc.

Chen, Y. & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment in Higher Education, 28*, 71-88.

Carkenord, D. M., & Stephens, M. G. (1994). Understanding student judgments of teaching effectiveness: A "policy capturing" approach. *Journal of Psychology, 128*(6), 675-682.

Deane, K., Ellis Hill, C., Dekker, K., Davies, P., & Clarke, C. E. (2003). A Delphi survey of best practice occupational therapy for Parkinson's disease in the United Kingdom. *British Journal of Occupational Therapy, 66*, 247-254.

Delucchi, M., & Pelowski, S. (2000). Liking or learning? The effect of instructor likeability and student perceptions of learning on overall ratings of teaching ability. *Radical Pedagogy, 2*(2), NP.

Epting, L. K., Zinn, T. E., Buskist, C., & Buskist, W. (2004)./ Students perspectives on the distinction between ideal and typical teachers. *Teaching of Psychology, 31*, 181-183.

Feeley, T. H. (2002). Evidence of halo effects in student evaluations of communication instruction. *Communication Education, 51(3),* 225-236.

Feldman, K. A. (1976). The superior college teacher from the student's view. *Research in Higher Education, 5*, 243-288.

Gaskin, C. J., O'Brien, A. P., & Hardy, D. J. (2003). The development of a professional practice audit questionnaire for mental health nursing in Aotearoa/New Zealand. *International Journal of Mental Health Nursing, 12*, 259-270.

Gordon, T. J., & Helmer, O. (1964). Report on a long range forecasting study. Rand Paper P-2982, Santa Monica, CA: Rand Corportation.

Harrison, P. D., Ryan, J. M., & Moore, P. S. (1996). College students' self-insight and common implicit theories in rating of teaching effectiveness. *Journal of Educational Psychology, 88*(4), 775-782.

Helmer, O. H., & Rescher, N. (1958). On the epistemology of the inexact sciences. Santa Monica, CA: Rand Corporation.

Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999). The dimensions of student's perceptions of teaching effectiveness. *Educational and Psychological Measurement, 59*(4), 580-596.

James, P., Aitken, P., & Burns, T. (2001). Research priorities for primary care mental health: A Delphi exercise. *Primary Care Psychiatry, 8*, 27-30.

Jenkins, D. A. (1996). A reflecting team approach to family therapy: A Delphi study. *Journal of Marital and Family Therapy, 22*, 219-238.

Jenkins, D. A., & Smith, T. E. (1994). Applying Delphi methodology in family therapy research. *Contemporary Family Therapy: An International Journal, 16*, 411-430.

Kessler, M., Nelson, B. S., Jurich, A. P., & White, M. B. (2004). Clinical decision-making strategies of marriage and family therapists in the treatment of adult childhood sexual abuse survivors. *American Journal of Family Therapy, 32*, 1-10.

Kher, N., Molstad, S., & Donahue, R. (1999). Using humor in the college classroom to enhance teaching effectiveness in 'dread courses'. *College Student Journal, 33(3),* 400-406.

Linstone, H. A., & Turoff, M. (2002). *The Delphi Method: Techniques and applications*. http://www.is.njit.edu/pubs/delphibook/ch4a.html

Lowman, J., & Mathie, V. A. (1993). What should graduate teaching assistants know about teaching? *Teaching of Psychology, 20*(2), 84-88.

Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal, 38(1),* 183-212.

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*(11), 1187-1197.

McKeachie, W. J., Lin, Y., Moffett, M. M., & Daugherty, M. (1978). Effective teaching: Facilitative vs. directive style. *Teaching of Psychology, 5(4),*193-194.

Meuleners, L. B., Binns, C. W., Lee, A. H., & Lower, A. (2002). Perceptions of the quality of life for the adolescent with a chronic illness by teachers, parents and health professionals: A Delphi study. *Child: Care, Health and Development, 28*, 341-349.

Miller, J. E., Dzindolet, M. T., Weinstein, L., Xie, X., & Stones, C. R. (2001). Faculty and students views of teaching effectiveness in the United States, China, and South Africa. *Teaching of Psychology, 28*(2), 138-142.

Moldrup, C., Morgall, J. M., & Almarsdottir, A. B. (2002). Perceived risk of future drugs--A Danish citizen Delphi. *Health, Risk and Society, 4*, 5-17.

Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports, 78*, 1363-1372.

Nussbaum, J. F. (1992). Effective teacher behaviors. *Communication Education, 41*, 167-180.

Perkins, D., Schenk, T. A., Stephan, L., & Vrungos, S. (1995). Effects of rapport, intellectual excitement, and learning on students' perceived ratings of college instructors. *Psychological Reports, 76*(2), 627-635.

Radmacher, S. A. & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *Journal of Psychology: Interdisciplinary and Applied, 135(3),* 259-268.

Schaeffer, G., Epting, K., Zinn, T., & Buskit, W. (2003). Student and faculty perceptions of effective teaching: A successful replication. *Teaching of Psychology, 30*(2), 133-136.

Sheehan, E. P. & DuPrey, T. (1999). Student evaluations of university teaching. *Journal of Instructional Psychology, 26(3),* 188-193.

Sinai, J., Tiberius, R. G., de Groot, J., Brunet, A., & Voore, P. (2001). Developing a training program to improve supervisor-resident relationships, Step 1: Defining the types of issues. *Teaching and Learning in Medicine, 13(2),* 80-85.

Spangenberg, H. H., & Theron, C. C. (2002). Development of a uniquely South African leadership questionnaire. *South African Journal of Psychology, 32*, 9-25.

Spencer, K. J., & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment and Evaluation in Higher Education, 27*, 397-409.

Tang, T. L. (1997). Teaching evaluation at a public institution of higher education: Factors related to the overall teaching effectiveness. *Public Personnel Management, 26*(3), 379-389.

Waters, M., Kemp, E., & Pucci, A. (1988). High and low faculty evaluations: Descriptions by students. *Teaching of Psychology, 15*, 203-204.

Author note

Correspondence can be addressed to Otto H. MacLin, Department of Psychology, University of Northern Iowa, Cedar Falls, IA 50614, or via Email at otto.maclin@uni.edu.