INFORMATION SCIENCE PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITB11767

This chapter appears in the book, *Selected Styles in Web-Based Research* edited by
Bruce L. Mann © 2006, Idea Group Inc.

Chapter 28

Correlational Research and the Internet

Bruce L. Mann
Memorial University, Canada

Learning Objectives

1. Define "correlational research."
2. Explain instances where percentiles or even raw data will do the job as well as a correlation.
3. Summarize the rationale for using peer assessment in Web-based educational research.
4. Describe the findings from recent applications of the Post And Vote Model applied to:
   a. Homework Web sites
   b. Verbalizations on Explorer Videos
5. Compare recent theses in Web-based teaching and learning that used correlation statistics.
6. Describe the identifying characteristics of Web-based correlational research.
7. Demonstrate the procedure for conducting Web-based correlational research.

Abstract

This chapter will discuss correlation of online data as a statistical technique to show significance and strength of agreement, and those times when percentiles or even raw data will do the job as well.

Correlation Research: What It Is

Correlation is a statistical technique that can show how strongly pairs of variables relate to one another, or agree on some measure. The most common type of correlation is called the Pearson Product Moment Correlation, named after Karl Pearson who developed this method to conduct agricultural research. The product moment part of the name comes from the way in which it is calculated, by summing up the products of the deviations of the scores from the average. The main result of a correlation is called the correlation coefficient (or "r"), and ranges from -1.0 to +1.0. The closer *r* is to +1 or -1, the more closely the two variables are related. While correlation coefficients are normally reported as r = (a value between -1 and +1), squaring them makes them easier to understand. The square of the coefficient (or r square) is equal to the percent of the variation in one variable that is related to the variation in the other. After squaring r, ignore the decimal point. An *r* of .5 means 25% of the variation is related (.5 squared =.25). An *r* value of .7 means 49% of the variance is related (.7 squared = .49).

A correlation report can also show another result of each test, namely: statistical significance. In this case, the significance level will tell you how likely it is that the correlations reported may be due to chance in the form of random sampling error. If you are working with small sample sizes, choose a report format that includes the significance level. This format also reports the sample size.

Correlation as a Test of Statistical Significance

A correlation report can also show another result of each test, namely: statistical significance. In this case, the significance level will tell you how likely it is that the correlations reported may be due to chance in the form of random sampling error. If you are working with small sample sizes, choose a report format that includes the significance level. This format also reports the sample size.

Reliability Coefficients

The reliability coefficient is a common form of correlation coefficient. Reliability is used to measure the extent to which an item, scale, or instrument will yield the same score when administered in different times, locations, or populations, when the two administrations do not differ in relevant variables.

Inter-Rater Reliability

As a method of measuring the agreement among coders in their analysis of unstructured data, inter-rater reliability is more closely aligned with experimental than qualitative methodologies. In this context, inter-rater reliability is used to establish consensus on the use of an instrument. For categorical data, consensus is measured as number of agreements divided by total number of observations. For continuous data, consensus is measured as the Pearson Product Moment correlation between the ratings for pairs of raters.

In practice, experimental researchers tend to report one of four different methods of determining inter-rater reliability, either: Cohen's Kappa, Kendall's Coefficient of Concordance, the intra-class correlation, or the Pearson Product Moment correlation (Huck, 2004, p. 84). Rourke, Anderson, Garrison and Archer (2001), for example, make an argument for using the Cohen's Kappa chance-corrected measure of inter-rater reliability over other statistics. Cohen's Kappa would be a good choice for analysis of the nominal data, where raters would have to classify each bulletin board discussion into one of five categories of online discussion. In some situations however, the judges' ratings are all raw scores, say from 0 to 30, and therefore best analyzed by the Pearson Product Moment correlation, the most frequently used correlational procedure in educational research (Cohen et al., 2000; Huck, 2004).

Online Peer Assessment

Peer assessment is the process where students evaluate one another's work (Topping, Smith, Swanson, & Elliot, 2000). Peer assessment requires that students exchange assignments, discuss responses, and rate one another's work using scoring systems or rating scales devised by the teacher (Airasian, 2001). Table 1 shows that different kinds of student-developed verbalizations and Web-based products can be peer-assessed online, and later correlated with the instructor's independent assessment.

Homework Web sites: In this ex post facto research, data from two sections of an undergraduate (pre-service teacher education) course were required to build an educational Web site, also known as a homework Web site. Working from an instructor-made rubric, each student used a rubric to develop a paper-based mock-up for an educational Web site, then a five-page online prototype from the mock-up, followed by four peer assessments of other student-made Web sites. Student-assessors were made "accountable," told in advance that the instructor would grade the quality of their peer assessments, and that this grade would be counted toward their final mark. Results in table 1 shows that the computed average of grades and justifying comments assigned by students were highly correlated, and statistically significant with grades assigned independently by the course instructor, assuring that the reliability of the research was high.

Verbalizations on explorer videos: Data from a different course was analyzed for inter-rater agreement. This time, in addition to developing a five-page online prototype from their paper mock-up of a Web site, each student in this course also had to use an Explorer

Table 1. Correlations of instructor and average of student peer assessments in large and small classes, showing the strengths of association and statistical significance (2-tailed)

Type / Sample	Level	Criterion	Agreement and shared variance	Sig.
Homework Web site Fall 04 Sec02, n=39	Pre-service	Instructor's rubric	$r = 0.385$* $r^2 = 0.148$	$p.= .015$
Homework Web site Fall 04 Sec04, n=39	Pre-service	Instructor's rubric	$r = 0.745$** $r^2 = 0.555$	$p.= .000$
Verbalizations of Web site Dev't Fall 02, n=66	Pre-service	Instructor's rubric.	$r = 0.701$** $r^2 = 0.491$	$p.= .000$
Discussion postings Winter 04, n=14	Masters, Software Development	Henri's (1992) model	r = 0.850** $r^2 = 0.7225$	$p.= .010$

Center to explain how they developed the Web site, then peer assess verbalizations of other students' Explorer videos (Mann, 1998, 1997, 1996). Explorer Centers were video analysis suites wherein a computer with a microphone was linked to a videotape recorder by a thin wire through a conversion box. Explorer Centres: (1) Modeled the appropriate learning behavior on a demo tape; (2) Gave each pre-service teacher a platform for generating the appropriate learning behavior on tape, and; (3) Provided a record from which to assess each pre-service teacher's verbalizations about the learning process. In course work involving individual Web-based tasks, Explorer Centers were found less intrusive than individual workstations and a means of collecting verbal protocols in the absence of the investigator's tape recorder. Aside from the watch-and-redo opportunities it afforded, the main difference between an Explorer Center and a regular computer workstation was the physical videotaped evidence of the teacher's achievement. The results in table 1 shows, as before, that the computed average of grades and justifying comments assigned by students were highly correlated, and statistically significant with grades assigned independently by the course instructor, assuring that the reliability of the research was high.

Despite the significant, positive correlations, however, new researchers should be aware that online peer assessment alone will likely not reduce errors in peer marking and comments due to "friendship marking," "collusive marking," "decibel marking" or "parasite marking" (Pond, Rehan, & Wade, 1995). Friendship marking is peer over-marking and a reluctance to provide critical comments. Collusive marking is a lack of differentiation within groups of markers. Decibel marking describes a context where individuals who dominate groups get the highest marks. Parasite marking describes a situation in which students fail to contribute but benefit from group marks. Anonymity and accountability can significantly reduce these problems (see the chapter by Wadhwa, Schulz, and Mann on "Effects of Anonymity and Accountability During Online Peer Assessment").

Thesis Research with Online Correlations

Knowledge, skill, and attitudes: David Taylor completed his Ed.D. in 2004 at Alliant International University In San Diego. His thesis was entitled, "Teaching Online: Postsecondary Faculty Perceptions of the Knowledge, Skills, and Attitudes Necessary to Succeed." Creswell's Concurrent Triangulation Strategy research design was used to collect, analyze, and compare quantitative and qualitative data from a Web-based survey. The survey instrument collected participant demographics, online teaching status and experience, the participant's perception of what knowledge, skill, and attitudes are necessary for teaching online, and finally a series of questions related to identifying the relative importance of knowledge, skill, and attitude items. The results were analyzed using a one-way ANOVA to determine statistical significance and Pearson's correlation to determine if there was any correlation between the responses and selected participant characteristics.

Interactivity and Web course quality: Barry Schlegel's Ed.D. research completed in 2003 at Rutgers University explored the relationship between interactivity in online distance education courses and course quality as measured by student satisfaction.

Online technologies and student success: In Joseph Corbeil's Ed.D. thesis completed in 2004 at the University of Houston, a correlational research design was used to study the relationships between three predictor variables and two criterion variables to test the hypotheses that, within a single group, these variables were statistically significantly related to one another. The purpose of the study was to describe the strength and direction of the relationship between online technologies self-efficacy, self-directed learning readiness, internal locus of control, and student success as measured by academic performance and student satisfaction.

Appropriate Applications of Correlation on the Web

Like all statistical techniques, a correlation is only appropriate for certain kinds of data. The Pearson correlation technique works best with linear relationships: As one variable gets larger, the other gets larger (or smaller) in direct proportion. It does not work well with curvilinear relationships (in which the relationship does not follow a straight line). An example of a curvilinear relationship is "age" and "use of e-mail." They are related, but the relationship doesn't follow a straight line. I suspect that people young and old alike ("age") are using e-mail now more than they used to—so there may be a small or no relationship of "age" to "e-mail use."

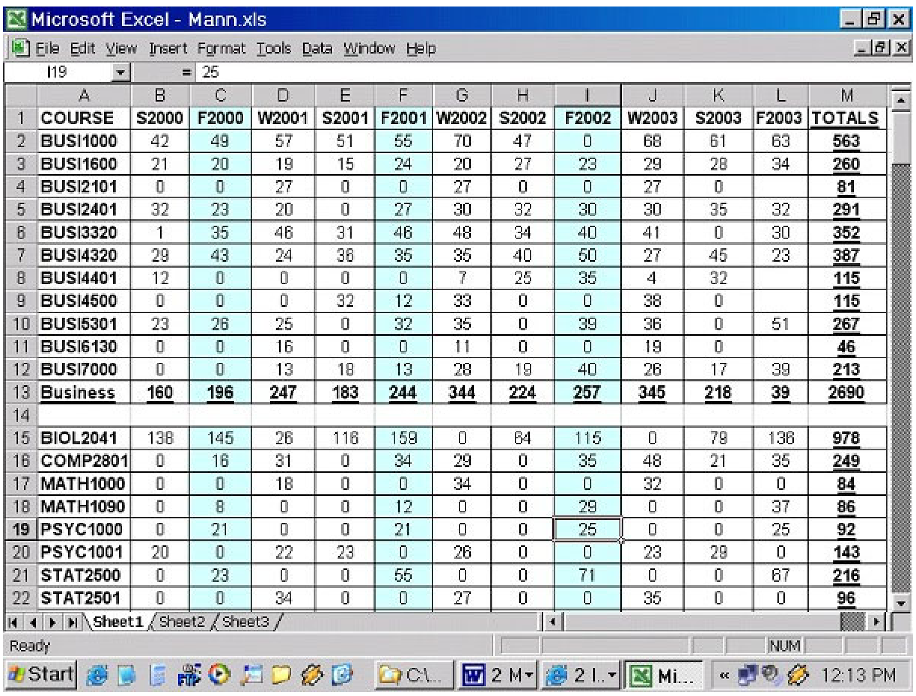
It's a good idea to keep in mind that the meaningfulness of a correlation is just as important as its truth. Educational researchers of Web-based treatments should be aware when making assumptions about correlations that changes occurring in two variables may not indicate a meaningful correlation between them. Take for example two likely changes: (1) "the number of student postings to a discussion board in your Web course have increased in the last few weeks" and (2) "mid-term grades have increased across the university from last year." It might be tempting to say that increases in "student postings in your course" and "mid-term grades in other courses" are somehow "correlated," which is likely a false assumption. In this example, the changes may be due to many other factors. The point being that change in one variable does not always directly relate to changes in another variable.

When Raw Data or Percentiles Will Do

Sometimes percentiles or even raw data will do the job as well as a correlation. Suppose you wanted to determine in general, just how many instructors at your university were using the Web exclusively in their teaching. To answer this question you need only ask the Manager of the distance education service at your institution for the raw data—a list of courses offered by the university on Web, as shown in Figure 1.

Figure 1 only shows us the "big picture". Figure 1 is showing us only a small portion of a very large data set – a lot of numbers displayed in a large spreadsheet. As is, it would

Figure 1. Spreadsheet screen capture showing number of courses-by-discipline each semester offered entirely on the Web



Microsoft Excel - Mann.xls

File Edit View Insert Format Tools Data Window Help

I19 = 25

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	COURSE	S2000	F2000	W2001	S2001	F2001	W2002	S2002	F2002	W2003	S2003	F2003	TOTALS
2	BUSI1000	42	49	57	51	55	70	47	0	68	61	63	563
3	BUSI1600	21	20	19	15	24	20	27	23	29	28	34	260
4	BUSI2101	0	0	27	0	0	27	0	0	27	0		81
5	BUSI2401	32	23	20	0	27	30	32	30	30	35	32	291
6	BUSI3320	1	35	46	31	46	48	34	40	41	0	30	352
7	BUSI4320	29	43	24	36	35	35	40	50	27	45	23	387
8	BUSI4401	12	0	0	0	0	7	25	35	4	32		115
9	BUSI4500	0	0	0	32	12	33	0	0	38	0		115
10	BUSI5301	23	26	25	0	32	35	0	39	36	0	51	267
11	BUSI6130	0	0	16	0	0	11	0	0	19	0		46
12	BUSI7000	0	0	13	18	13	28	19	40	26	17	39	213
13	Business	160	196	247	183	244	344	224	257	345	218	39	2690
14													
15	BIOL2041	138	145	26	116	159	0	64	115	0	79	136	978
16	COMP2801	0	16	31	0	34	29	0	35	48	21	35	249
17	MATH1000	0	0	18	0	0	34	0	0	32	0	0	84
18	MATH1090	0	8	0	0	12	0	0	29	0	0	37	86
19	PSYC1000	0	21	0	0	21	0	0	25	0	0	25	92
20	PSYC1001	20	0	22	23	0	26	0	0	23	29	0	143
21	STAT2500	0	23	0	0	55	0	0	71	0	0	67	216
22	STAT2501	0	0	34	0	0	27	0	0	35	0	0	96

Sheet1 Sheet2 Sheet3

Ready NUM

Start C:\... 2 M 2 I... Mi... 12:13 PM

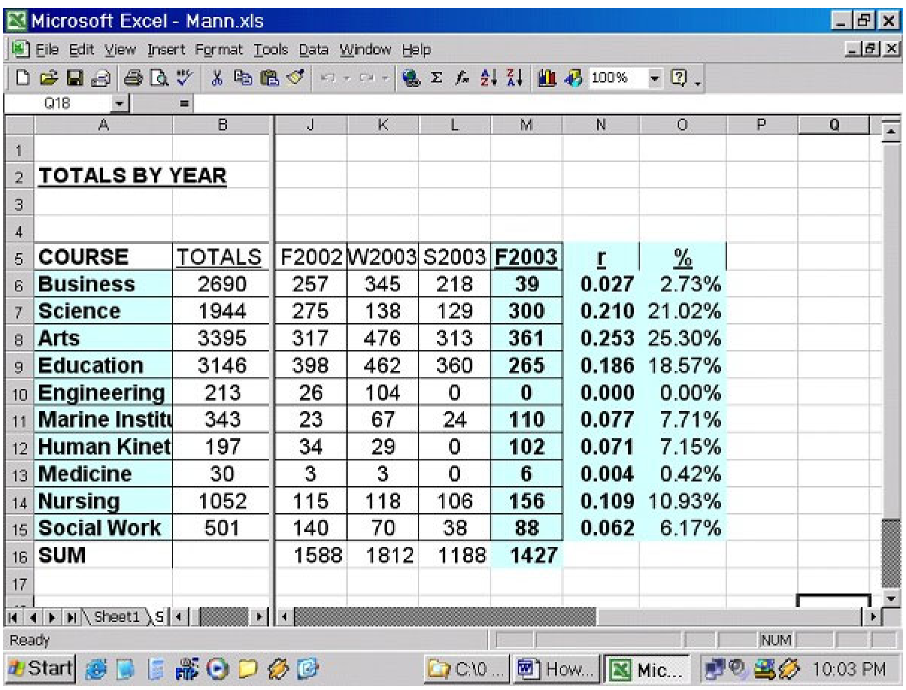
be difficult to make sense of anything in this form. Some data massaging is needed, and another question, to prompt our analysis. New researchers should also be aware at this point, that these data will not show us different students taking courses in any given semester, say Arts students versus Science students, because a single student may take both Arts and Science courses, and therefore might be represented more than once in these data.

Let's say you wanted to compare disciplines against one another to determine their Web course development activity. You could copy the summary data from each semester for each discipline in your large spreadsheet onto a separate sheet, and display the numbers as either raw scores, "r" correlations, or percentages of the total number of course offered, as shown in Figure 2.

Now, suppose you believed in the possibility, and wanted to confirm that there was a growing trend at your university toward Web-based education. A line graph would clearly show a rise or fall over the semesters for each discipline. The line graph in Figure 3 generated by the spreadsheet program indicates there has been a steady rise in Business and Nursing Web courses at the university, while Science is generally showing a fall in courses developed entirely for Web over these 10 semesters.

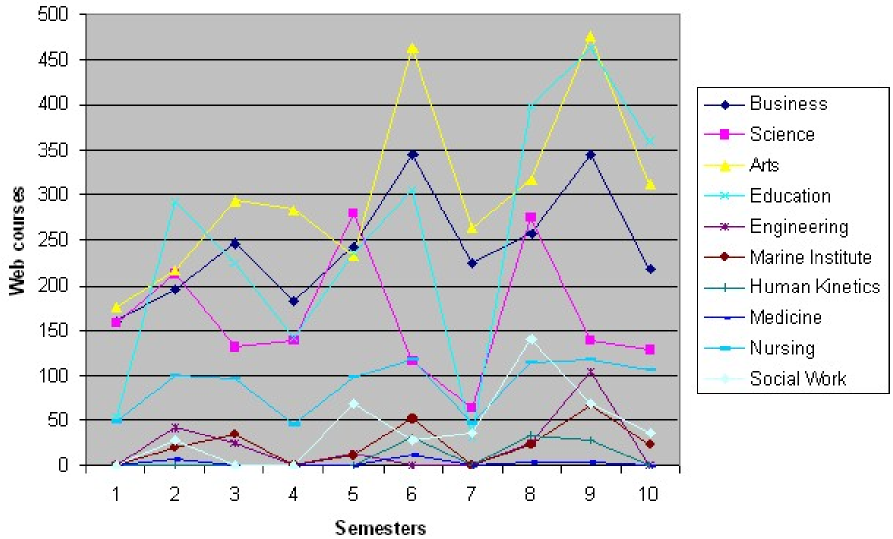
Lastly, let's say that you wanted to find out how the Web was being used by instructors, you could use a technology adoption model, such as CBAM or the Phase Theory of Web

Figure 2. Spreadsheet screen capture showing number of courses-by-discipline each semester offered entirely on the Web



COURSE	TOTALS	F2002	W2003	S2003	F2003	r	%
Business	2690	257	345	218	39	0.027	2.73%
Science	1944	275	138	129	300	0.210	21.02%
Arts	3395	317	476	313	361	0.253	25.30%
Education	3146	398	462	360	265	0.186	18.57%
Engineering	213	26	104	0	0	0.000	0.00%
Marine Institu	343	23	67	24	110	0.077	7.71%
Human Kinet	197	34	29	0	102	0.071	7.15%
Medicine	30	3	3	0	6	0.004	0.42%
Nursing	1052	115	118	106	156	0.109	10.93%
Social Work	501	140	70	38	88	0.062	6.17%
SUM		1588	1812	1188	1427		

Figure 3. Number of university courses-by-discipline offered entirely on the Web over 10 semesters



Course Behavior, discussed in another chapter. To answer this question, you'd need to gather interview data with course instructors over an extended period of time.

Student Exercise

1. From the list of courses shown in the far-left column in Figure 1, how many business courses are being offered entirely over the Web?
2. From the data shown in Figure 2, rank the order of the number of courses offered in the Fall 2003 semester by discipline from 1st to 10th.
3. Calculate the correlations of Web course per discipline to total Web courses offered in Fall 2002, Winter 2002, and in Spring 2003.
4. From the data shown in Figure 2, summarize in your own words the trends of Web courses offered by Engineering and Education.

References

Airasian, P. W. (2001). *Classroom assessment: Concepts and applications* (4th ed.). New York: McGraw-Hill.

Anderson, A., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science*, *29*, 1-32.

Brown, K. (2002). *A case study describing online discussion in a tenth-grade biotechnology class.* Ph.D. University of Virginia. Dissertation Abstracts International (DAI-A 63/03, p. 909).

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). Taylor & Francis, Inc.

Corbeil, J. R. (2003). *Online technologies self-efficacy, self-directed learning readiness, and locus of control of learners in a graduate-level Web-based distance education program.* Ed.D., University of Houston, AAT 3094002.

Gall, M. D., Borg, W. R., & Gall, J. P. (2003). *Educational research: An introduction* (7th ed.). White Plains, New York: Longman.

Hara, N., Bonk, C., & Angeli, C. (2000). Content analysis of online discussions in an applied educational psychology course. *Instructional Science*, *28*(2), 115-152.

Henri, F. (1992). Computer conferencing and content analysis (pp. 117-136). In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing*. Berlin: Springer-Verlag.

Huck, S. W. (2004). *Reading statistics and research* (4th ed.). Boston: Pearson Education.

Mann, B. L. (1996). Preservice applications of Explorer Centers for general computing. In J. White (Ed.). *The technology and teacher education conference annual* 1996. Needham, MA: Allyn & Bacon.

Mann, B. L. (1997, September 17-19). Explorer centers (ECs): Good practice and assessment for teacher education. School education in the information society: *Proceedings of the Open Classroom II Conference,* Crete, Greece.

Mann, B. L. (1998, June). Instructional design for online learning: A case study of WebCT developers. Universities in a digital era: Transformation, Innovation and tradition. *Proceedings of The 7th Annual European Distance Education Network (EDEN) Conference,* University of Bologna, Italy.

Mann, B. L. (1999a). Web course management. *Australian Educational Computing*, *14*(1).

Mann, B. L. (1999b). Web course management in higher education. *Bulletin of The Canadian Society for the Study of Higher Education.*

Mann, B. L. (2000a). Phase theory: A teleological taxonomy of Web course management (pp. 3-26). In B. L. Mann (Ed.), *Perspectives in Web course management.* Toronto, ON: Canadian Scholar's Press.

Mann, B. L. (2000b). WebCT: serving educators in newfoundland & labrador. *The New Morning Watch: Educational and Social Analysis*, *28*,1/2.

Mann, B. L. (2005). The post and vote model of web-based peer assessment: Testing the agreement between instructor and students. In M. Hricko S. Howell, and D. Williams (Eds.), *Online assessment and measurement* (Vol.3) Hershey, PA: IGP.

Mann, B. L., Whitmore, M., & Bussey, W. (2000). Network-intensive research at Memorial University. *The New Morning Watch: Educational and Social Analysis, 28*,1/2.

Mertler, C., & Charles, C. (2005). *Introduction to research* (5th ed.). New York: Pearson Education Inc.

Pond, K., Rehan, U., & Wade, W. (1995). Peer review: A precursor to peer assessment. *Innovations in Education and Training International*, *32*(4), 314-323.

Rourke, L., Anderson, T., Garrison, D., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education, 12.*

Rose, M. (2002). *Cognitive dialogue, interaction patterns, and perceptions of graduate students in an online conferencing environment under collaborative and cooperative structures.* Ed.D., Indiana University. Dissertation Abstracts International (DAI-A 63/04, p. 1249).

Rushton, C., Ramsey, P., and Rada, R. (1993). Peer assessment in a collaborative hypermedia environment: A case study. *Journal of Computer-Based Instruction*, *20*(3), 75-80.

Schlegel, B. (2003). *Relationship between interactivity in online distance education courses and course quality as measured by student satisfaction.* Doctoral dissertation, Rutgers University.

Schuttler, R., & Burdick, J. (in press). Creating a Unified System of Assessment. In M. Hricko, S. Howell, and D. Williams (Eds.), *Online assessment and measurement* (Volume One). Hershey, PA: Idea Group Publishing.

Taylor, D. (2004). *Teaching online: Postsecondary faculty perceptions of the knowledge, skills, and attitudes necessary to succeed.* Doctoral dissertation, Alliant International University in San Diego.

Topping, K.J., Smith E.F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education*, *25*(2), 149-169.

Wadhwa, G., Mann, B. L., Schulz, H. (2004). *Effects of anonymity and peer-accountability during peer assessment in a graduate Web-based education research methods course.* Paper Session on Teaching Educational Research, Measurement and Research Methodology, Division D, American Educational Research Association. San Diego, CA.