

New Measures for Course Evaluation in Higher Education and their Relationships with Student Learning

Theodore W. Frick

Rajat Chadha

Carol Watson

Emilija Zlatkowska

School of Education

Indiana University Bloomington

Paper presented at the annual meeting of the

American Educational Research Association

Denver, Colorado

May 1, 2010

Abstract

The authors have developed a new instrument to measure teaching and learning quality (TALQ) in postsecondary education. In a study of 464 students in 12 courses, when students agreed that their instructors used First Principles of Instruction and also agreed that they experienced academic learning time (ALT), they were about 5 times more likely to achieve *high levels* of mastery of course objectives and 26 times less likely to achieve *low levels* of mastery, according to independent instructor assessments. First Principles of Instruction include: authentic tasks for students to learn to do, activation of student learning, demonstration of what is to be learned, student application with instructor feedback, and student integration of what they have learned into their personal lives.

Problem

Few items typically used in traditional course evaluations are empirically associated with student learning achievement. Among the few items which have a consistent positive relationship with learning are student ratings on global items such as “This was an outstanding course” or “The instructor of this course was outstanding”. Even so, such ratings correlate only moderately with student achievement (averages of 0.47 and 0.43, respectively—cf., Cohen, 1981; Feldman, 1989; Kulik, 2001). Nonetheless, ratings on such global items do not suggest ways to improve a course or how it is taught. Other more specific items may indicate needed improvements, but there is a paucity of empirical evidence relating these ratings to student learning achievement—see reviews of literature by Frick, et al. (2009, 2010b).

Objectives and Theoretical Framework

Frick, et al. (2009, 2010a, 2010b) have developed a new course evaluation instrument for assessing Teaching and Learning Quality (TALQ). While the initial TALQ instrument has been slightly modified based on reliability analyses from earlier studies, what is noteworthy about the TALQ is that *a priori* scales have been constructed according to instructional theories and other important variables which have been empirically associated with student learning achievement. In particular, new scales were developed for student ratings of First Principles of Instruction (Merrill, 2002; Merrill, Barclay & van Schaak, 2008) and for student rating of his or her own Academic Learning Time (ALT) (cf. Rangel & Berliner, 2007). In addition, TALQ rating scales are included that are consistent with Cohen’s (1981) meta-analysis that pertain to global course/instructor quality and student learning progress, both empirically demonstrated in meta-analyses to be positively correlated with student achievement at the college level.

Academic learning time (ALT) refers to repeated successful student engagement in learning activities that are relevant to curriculum goals. Rangel and Berliner (2007) define ALT as “...the amount of time that students spend on rigorous tasks at the appropriate level of difficulty for them” (p. 1) when those students are “... engaged in tasks relevant to curriculum expectations and assessments” (p. 1). That is, those tasks need to be in a student’s zone of proximal development (Vygotsky, 1978), which means that the tasks cannot be done by a student alone but can with assistance or guidance, and that those tasks are sequenced to lead towards curriculum goals, not just repeatedly doing the same tasks successfully. Numerous studies have found significant positive correlations between ALT and student achievement (cf. Kuh *et al.*, 2007; Berliner, 1990; Brown & Saks, 1986).

First Principles of Instruction are relevant to complex learning of authentic, real-world, whole tasks. Based on a synthesis of instructional design theories, Merrill (2002) claimed that student learning will be promoted when: 1) students are expected to learn to do authentic (real-world) tasks, 2) student learning is activated by connecting what they already know or can do with what is to be newly learned, 3)

students are exposed to demonstrations of what they are to learn, 4) they have opportunities to try out what they have learned with instructor scaffolding and feedback, and 5) they integrate what they have learned into their personal lives. If one or more of these First Principles are missing during instruction, Merrill argues that learning will be negatively impacted.

Solving *authentic* problems or performing *authentic* learning tasks is also fundamental for complex learning in the 4C/ID model of instructional design (van Merriënboer & Kirschner, 2007). Like Merrill (2002), they discuss the use of real-world or authentic tasks as being central to the learning process. They recommend grouping whole tasks into classes, where those task classes are arranged from simple to complex in terms of what is required for successful performance. Within each task class, there is not only repetition but also variation of tasks. As whole tasks are repeated with variation within each task class, teacher scaffolding of student performance (e.g., assistance, feedback, coaching) is gradually withdrawn until learners can successfully perform whole tasks in that task class on their own. Then this cycle is repeated for the next more-complex, whole-task class.

When students engage in *authentic* whole tasks, they are provided with opportunities to make connections between knowing *that-one*, knowing *how* and knowing *that* (cf. Frick, 1997; Maccia, 1987). Moreover, when students engage in these tasks willingly, such connections are strengthened by their emotional involvement and sense of purpose as they directly experience objects of learning (cf., Greenspan & Benderly, 1997; Dewey, 1916; Montessori, 1965). Estep (2003) emphasizes the significance of immediate awareness of objects of experience for knowing how:

Knowing how is far more fundamental in our intelligence than *knowledge that* because it is logically, epistemologically, and temporally *prior* to our knowing propositional (knowledge *that*) statements.... In part we *know how* to do these things because of the immediate awareness ... found in the patterns of our actions, interactions and transactions with objects around us and in us... [O]ur immediate awareness ... extends beyond our language, beyond our ability to classify those objects of our immediate awareness. (pp. xvii-xviii, italics original)

Objective and Research Questions

Only one empirical study had been conducted to verify Merrill's (2002) claim that First Principles promote student learning. The present study sought to do so in the context of postsecondary education.

Research questions addressed:

1. What are the relationships among the TALQ scales and student learning achievement, as indicated by independent instructor assessments of student mastery of course objectives?
2. What are the odds that students are rated as high and low masters of course objectives by their instructors, when comparing students who do and do not agree that both First Principles and ALT occurred?

Method

Participant selection

In collaboration with staff from a teaching center at a large Midwestern university, a recruitment e-mail was sent to faculty that sought volunteers who were willing to have the TALQ instrument used in their classes, in addition to normal course evaluations. Researchers met individually in advance with each participating faculty member to explain the purpose of the study, obtain his or her signed consent to participate in the study, and to arrange a specific class day and time near the end of the semester when researchers could administer the TALQ.

Respondents

Data were collected from 464 students in 12 different courses taught by 8 instructors in business, philosophy, history, kinesiology, social work, informatics, nursing, and health, physical education and recreation. The number of student respondents who completed the TALQ ranged from 16 to 104 in the 12 classes, though in 10 of the 12 classes the range was from 22 to 53.

Instrument

The first page of the current version of the TALQ instrument included items on gender, expected grade, student year in school, and self-reported mastery of course objectives. Subsequent pages on the survey instrument included 40 randomly ordered items that attempted to measure the TALQ scales via student ratings. Students did not know which items belonged to each scale. Each item was rated on a Likert scale (1 = strongly disagree; 2 = agree; 3 = undecided; 4 = agree; 5 = strongly agree). Instructors independently rated student mastery on a 10-point scale (1 = nonmastery ... 10 = mastery). These TALQ scales and Cronbach α coefficients of reliability (internal consistency) are listed below:

Authentic Problems Scale: First Principles of Instruction: $\alpha = 0.690$

- I performed a series of increasingly complex authentic tasks in this course.
- I solved authentic problems or completed authentic tasks in this course.
- In this course I solved a variety of authentic problems that were organized from simple to complex.

Activation Scale: First Principles of Instruction: $\alpha = 0.812$

- I engaged in experiences that subsequently helped me learn ideas or skills that were new and unfamiliar to me.
- In this course I was able to recall, describe or apply my past experience so that I could connect it with what I was expected to learn.
- My instructor provided a learning structure that helped me to mentally organize new knowledge and skills.
- In this course I was able to connect my past experience to new ideas and skills I was learning.
- In this course I was not able to draw upon my past experience nor relate it to new things I was learning. (reverse-coded)

Demonstration Scale: First Principles of Instruction: $\alpha = 0.830$

- My instructor demonstrated skills I was expected to learn in this course.
- Media used in this course (texts, illustrations, graphics, audio, video, computers) were helpful in learning.
- My instructor gave examples and counter-examples of concepts that I was expected to learn.
- My instructor did not demonstrate skills I was expected to learn. (reverse-coded)
- My instructor provided alternative ways of understanding the same ideas or skills.

Application Scale: First Principles of Instruction: $\alpha = 0.758$

- My instructor detected and corrected errors I was making when solving problems, doing learning tasks, or completing assignments.
- I had opportunities to practice or try out what I learned in this course.

- My course instructor gave me personal feedback or appropriate coaching on what I was trying to learn.

Integration Scale: First Principles of Instruction: $\alpha = 0.780$

- I had opportunities in this course to explore how I could personally use what I learned.
- I see how I can apply what I learned in this course to real life situations.
- I was able to publicly demonstrate to others what I learned in this course.
- In this course, I was able to reflect on, discuss with others, and defend what I learned.

First Principles of Instruction — Combined Scale: $\alpha = 0.881$

An average scale score was computed for each student on each First Principles scale above. Then an overall First Principles scale score was computed by averaging these means for each student.

Academic Learning Time Scale: $\alpha = 0.763$

- I did not do very well on most tasks in this course, according to my instructor's judgment of the quality of my work. (reverse-coded)
- I frequently did very good work on projects, assignments, problems and/or activities for this course.
- I spent a lot of time doing tasks, projects and/or assignments, and my instructor judged my work of high quality.
- I put a great deal of effort and time into this course, and it has paid off—I believe that I have done very well overall.

Learning Progress Scale: $\alpha = 0.935$

- Compared to what I knew before I took this course, I learned a lot.
- I learned a lot in this course.
- I learned very little in this course. (reverse-coded)
- I did not learn much as a result of taking this course. (reverse-coded)

Satisfaction Scale: $\alpha = 0.926$

- I am very satisfied with how my instructor taught this class.
- I am dissatisfied with this course. (reverse-coded)
- This course was a waste of time and money. (reverse-coded).
- I am very satisfied with this course.

Global Quality Scale: $\alpha = 0.915$

- Overall, I would rate the quality of this course as outstanding.
- Overall, I would rate this instructor as outstanding.
- Overall, I would recommend this instructor to others.

Student Mastery of Course Objectives

Each instructor independently rated each student about one month after the course was over, using a 10-point mastery scale. Student mastery ratings were based on instructor evaluation

of their performance in the course—e.g., from scores on quizzes, exams, papers, projects, reports, presentations, class participation, assignments, etc.

Results

Relationships among Student Ratings on TALQ Scales and Instructor Ratings of Student Mastery

It can be seen from Table 1 that First Principles of Instruction ratings are positively and very highly correlated with Global Quality, Student Satisfaction, ALT, and Learning Progress (Spearman ρ 's ranging from 0.583 to 0.778, $p < 0.0005$). ALT is significantly correlated with Learning Progress ($\rho = 0.498$, $p < 0.0005$). The highest correlation with instructor rating of Student Mastery is Academic Learning Time ($\rho = 0.362$, $p < 0.0005$).

Table 1. Spearman Correlations among TALQ Scales

	First Principles	Global Quality	Student Satisfaction	ALT	Learning Progress	Student Mastery ^a
First Principles	1.000	0.774	0.778	0.583	0.725	0.115 ^b
Global Quality		1.000	0.848	0.528	0.664	0.180
Student Satisfaction			1.000	0.557	0.746	0.202
Academic Learning Time				1.000	0.498	0.362
Learning Progress					1.000	0.136 ^c
Student Mastery						1.000

a: 10-point scale used here for independent instructor ratings of student mastery of course objectives; b: $p = 0.014$; c: $p = 0.003$; all remaining correlations are significant at $p < 0.0005$; $n = 464$.

Pattern Analysis

We would expect students to be more motivated when instructors use First Principles of Instruction, because students are expected to solve authentic or real-world problems as well as to integrate what they have learned into their personal lives. In other words, what they learn is expected to be more relevant and meaningful (see Keller, 1987). If students are more highly motivated, then they would be expected to be engaged more often in learning tasks. Furthermore, if instructors demonstrate what students are expected to learn and also provide feedback and scaffolding when students themselves try, we would expect student engagement to be successful more often—i.e., more Academic Learning Time (ALT). The research on ALT indicates that the more frequently students are successfully engaged, the higher they tend to score on tests of learning achievement.

Analysis of Patterns in Time (APT) was used to further investigate these relationships (Frick, 1990). With the exception of the student mastery scale (recoded as low, medium and high), remaining scales were recoded for ‘agreement’ = ‘Yes’ if the scale score was greater than 3.5, and ‘agreement’ = ‘No’ if the student’s scale score was less than or equal to 3.5. The reasoning for this coding system was that on the original Likert scale, ‘agree’ was coded as ‘4’ and ‘strongly agree’ as ‘5’; thus, any mean scale score that was closer to ‘4’ or ‘5’ was interpreted as agreement with that scale; otherwise it was interpreted as *not* in agreement (strongly disagree = ‘1’, disagree = ‘2’, or undecided = ‘3’).

Table 2. Results for the APT Query: *If Agree that First Principles of Instruction occurred is ___ and Agree that ALT occurred is ___, then Instructor rating of student mastery is of course objectives is ___?*

		Agree that First Principles of Instruction occurred							
		No				Yes			
		Agree that ALT occurred				Agree that ALT occurred			
		No		Yes		No		Yes	
		Count	Column N %	Count	Column N %	Count	Column N %	Count	Column N %
Instructor rating of student mastery of course objectives	Low (0-5)	15	33.3%	1	6.2%	1	2.2%	2	1.3%
	Medium (6-8)	28	62.2%	10	62.5%	42	93.3%	114	76.0%
	High (8.5-10)	2	4.4%	5	31.2%	2	4.4%	34	22.7%
	Total	45	100.0%	16	100.0%	45	100.0%	150	100.0%

In Table 2, the APT Query addresses the combination of First Principles, ALT and student mastery ratings. It can be seen that for the APT Query, ‘If Agreement on First Principles is Yes *and* Agreement on Successful Engagement (ALT) is Yes, then Instructor Rating of Student Mastery is High?’ is true in 34 out of 150 cases, yielding a conditional probability estimate of 0.227. On the other hand, ‘If Agreement on First Principles is No *and* Agreement on Successful Engagement is No, then Instructor Rating of Student Mastery is High?’ is true in 2 out of 45 cases, yielding a conditional probability estimate of 0.044. Thus, a student is about 5.2 times as likely to be rated by his or her instructor (and himself or herself) as a high master of course objectives when that student agreed that First Principles occurred and also agreed that she or he experienced ALT (successful engagement), compared with *not* agreeing that First Principles and ALT occurred. The odds of 5.2 to 1 are computed as a ratio of the two probabilities: (0.227/0.044). The odds are about 25.6 to 1 of being a student being rated as a *low* master of course objectives by his or her instructor when that student did not agree that First Principles and ALT occurred (0.333/0.013 = 25.6), compared with being rated as a low master when a student agreed that both First Principles and ALT did occur.

One can also see in Table 2 that when a student agreed that First Principles occurred but *not* ALT, she or he is about 1.5 times more likely to be rated at a medium level of mastery, compared with agreeing that ALT occurred but *not* First Principles (0.933/0.625).

Note that a subset of 256 cases out of 464 was selected for this analysis in which the student's self-rating of his or her mastery level was exactly the same as the instructor's independent assessment of that student's mastery level. The rationale was that these were the cases in which the assessments of student mastery were the most accurate, since both the instructor and student independently agreed on that student's level of mastery of course objectives. Nonetheless, similar patterns as above occurred when APT was done on the entire sample.

Factor Analysis: Instructional Quality

Spearman correlations were generally very high among the scales related to the quality of the course and instructor (see Table 1). We further wondered: Are these scales measuring the same overall construct, perhaps something that might be called 'Instructional Quality?' To answer this question, we conducted a factor analysis of scales about a course and how it is taught—things over which *instructors* have direct control or which should result from their efforts, not those of the students.

Table 3. Loadings from factor analysis of student ratings of instructors and courses.

TALQ Scale	Loading
	Factor 1
Student satisfaction	.895
Global quality	.883
Demonstration	.815
Integration	.777
Authentic Problems	.740
Activation	.685
Application	.621

We excluded from this factor analysis scales related to aspects of a course and learning which an instructor cannot control—i.e., *student* learning progress, mastery of course objectives and academic learning time—since these are elements that are affected by *student* effort and engagement. A student's volition is under his or her own control, e.g., that student can *choose* to go to or skip class, participate or not in class, and do or ignore tasks related to course objectives.

We used the image analysis method of extraction, and factor loadings ranged from 0.895 to 0.621, as can be seen in Table 3. These are strong loadings and are consistent with the high correlations reported in Table 1. What is noteworthy is that all five First Principles of Instruction scales load on the same factor that includes student satisfaction with the course and instructor and student global ratings of the course and instructor quality. It should be noted that students did not know what scales we were measuring, since individual scale items were randomly distributed within the TALQ instrument. These seven TALQ scales measure a single factor that could be termed 'Instructional Quality.'

Discussion

Results from this study are consistent with well-established empirical evidence that supports the positive relationship between Academic Learning Time and student achievement (e.g., Kuh *et al.*, 2007; Rangel & Berliner, 2007). We found significant positive correlations between student ratings of their ALT and learning progress ($\rho = 0.498, p < 0.0005$), and between ALT and instructor ratings of student mastery of course objectives ($\rho = 0.362, p < 0.0005$). Moreover, there were strong positive correlations between ALT and student ratings of overall course and instructor quality ($\rho = 0.528, p < 0.0005$), and between ALT and First Principles of Instruction ($\rho = 0.583, p < 0.0005$).

In our study, we did not utilize common end-of-course exams in multi-section courses taught by different instructors as measures of student learning achievement, as was done in studies included in Cohen's (1981) meta-analysis. However, we did obtain independent instructor assessments of student mastery of course objectives on a 10-point scale. Instructors did not know how each student rated their course on the TALQ scales. Those instructor assessments were based on student performance—not only in class, but also by indicators of learning that those instructors normally used to evaluate their students. Those indicators of mastery of course objectives were specific to the goals of each course in a variety of disciplines that included business, informatics, health sciences, history, kinesiology, nursing, philosophy, and social work. Instructors in this study never saw individual student ratings on the TALQ and had no vested interest in the outcomes. We did not inform instructors or their students *in advance* about TALQ scales designed to measure First Principles of Instruction, whose items were randomly ordered within the survey instrument.

Although not reported in the results above, there was a high positive correlation between grades students expected to receive and independent mastery ratings from their instructors ($\rho = 0.584, p < 0.0005$). We were not able to gain access to student grades in the course due to privacy laws, but we did note that, from student self-reports on their expected grade, over 93 percent expected to receive an A or B in the course.

Indeed, ratings on high mastery of course objectives were more discriminating than course grades. Only 17 percent of the students were rated as high masters of course objectives by both those students and their instructors, while 49 percent of the students expected to receive an A in the course. Agreement between student self-ratings and instructor ratings of student mastery was statistically significant when corrected for chance agreement ($\kappa = 0.17, p < 0.0005$). Since there was some disagreement between ratings of mastery of course objectives, we restricted our pattern analyses to cases in which instructors and students both independently agreed on the student's mastery level (256 cases).

Results from Analysis of Patterns in Time in this study are consistent with theoretical predictions from Merrill (2002) on First Principles of Instruction. Merrill sought to identify general principles of instruction that were not specific to a particular instructional program or content domain, but rather those principles which were common to instructional design theories. He sought:

... to identify and articulate the prescriptive design principles on which these various [instructional] design theories and models are in essential agreement.... A *principle* ... is a relationship that is always true under appropriate conditions regardless of program or practice... A *practice* is a specific instructional activity. A *program* is an approach consisting of a set of prescribed practices. Practices always implement or fail to implement underlying principles whether these principles are specified or not. A given instructional approach may only emphasize the implementation of one or more of these instructional principles. The same principles can be implemented by a wide variety of programs and practices. (p. 43)

The results of this study are also consistent with the *Theory of Immediate Awareness* (Estep, 2003; 2006). Estep discusses the “intelligence of doing” as well as findings from neuroscience that support the necessity for immediate awareness (knowing the unique) for coming to *know how*. Immediate awareness is the relation between the learner and *sui generis* objects:

Because *sui generis* objects in the immediate awareness relation are not class objects, they are not linguistic objects either. As such, they cannot be reduced in any way to objects of *knowledge by description*. The immediate awareness knowing of such objects is *knowing the unique....* These objects are very real to subjects [learners]; they are in immediate relations with subjects and have a direct affect upon their intentional and intelligent behavior. (Estep, 2006, p. 209)

Greenspan (1997) has arrived at the same conclusion through considerable clinical and neurological evidence. Greenspan’s findings are consistent with Estep (2006). Their conclusions contradict the long-held notions of separating affect, cognition and behavior—e.g., as indicated in Bloom, *et al.*’s well-known Taxonomy of Educational Objectives (cf. Krathwohl, 2002): cognitive, affective and psycho-motor domains. Greenspan argues that this Western tradition, based on how ancient Greeks characterized mind, has literally blinded us to the central role of affect in organizing our experience:

if ... information is dual-coded according to its affective and sensory qualities, then we have a structure or circuitry set up in our minds that enables us to retrieve it readily Affects enable us to identify phenomena and objects and to comprehend their function and meaning. Over time, they allow us to form abstract notions of interrelations.... Affect, behavior and thought must be seen as inextricable components of intelligence. For action or thought to have meaning, it must be guided by intent or desire (i.e., affect). Without affect, both behavior and symbols have no meaning. (Greenspan, 1997, pp. 30-37)

Research from neuroscience and clinical experience that supports Estep’s *Theory of Immediate Awareness* and Greenspan’s conclusions leaves little doubt as to the vital importance of *authentic experience*—i.e., through unmediated sensory interaction with the real world—in human learning and growth of the human mind. These findings are consistent with Merrill’s principles for engaging students in solving *real-world* problems and performing *authentic* tasks (Principle 1) and with integration of what is learned into student’s own lives (Principle 5).

Conclusion

Although Merrill et al. (2008) stated that the real value of the First Principles is in the design of instruction, they also argued that “learning from a given program will be facilitated in direct proportion to its implementation of these principles” (p. 175). Indeed, this was born out in our study. While academic learning time (ALT) is under the control of the student, use of the First Principles of Instruction in a classroom is something that instructors can control.

On a typical course evaluation, low scores on global items or low scores on student satisfaction *do not tell instructors anything about how to improve their teaching in ways that are likely to also improve student mastery of course objectives*. On the other hand, the TALQ scales on the First Principles of Instruction can be used to identify areas in which teaching and course design can be improved.

For example, a course can be modified to become task-centered—i.e., be structured around a series of increasing complex authentic problems—in contrast with a topic-centered structure (cf., Merrill, 2007). Activation of student learning can be strengthened by having students directly or vicariously experience objects of learning. Demonstrations of what is to be learned can be supplemented with multimedia examples (e.g., via video podcasts). Instead of having students passively listen to lectures, class time can be spent having students work in teams on authentic tasks while instructors can provide

immediate feedback, scaffolding or coaching to those teams. Integration activities can also be included in a course where students are provided with opportunities to reflect upon, publically demonstrate and discuss what they have learned.

If instructors make these kinds of changes in their courses, such changes would be expected to result in higher TALQ scale scores in their course evaluations. Most importantly, students would be expected to increase their ALT—i.e., successful engagement in tasks relevant to course objectives. Along with increased ALT, increased student learning achievement and satisfaction with the course and instructor would be expected. Finally, global ratings of items traditionally used to evaluate the overall quality of courses and instruction would likewise be expected to increase.

Finally, Frick et al. (2009, 2010b) empirically found that relationships among TALQ scales were consistent across a very wide range of subject matter in multiple disciplines, in face-to-face and in online learning contexts, and in undergraduate and graduate level courses. Similar patterns were found in the present study. Use of TALQ scales in end-of-course evaluations would be an efficient way to measure improvement in the quality of instruction and learning in postsecondary education.

References

- Berliner, D. (1990). What's all the fuss about instructional time? In M. Ben-Peretz & R. Bromme (Eds.), *The nature of time in schools: Theoretical concepts, practitioner perceptions*. New York: Teachers College Press.
- Brown, B. & Saks, D. (1986). Measuring the effects of instructional time on student learning: Evidence from the Beginning Teacher Evaluation study. *American Journal of Education*, 94(4), 480-500.
- Cohen, P. (1981). Student ratings of instruction and student achievement. A meta-analysis of multisection validity studies. *Review of Educational Research*, 51(3), 281-309.
- Dewey, J. (1916). *Democracy and education*. New York: The Free Press.
- Estep, M. (2003). *A theory of immediate awareness: Self-organization and adaptation in natural intelligence*. Boston: Kluwer Academic Publishers.
- Estep, M. (2006). *Self-organizing natural intelligence: Issues of knowing, meaning and complexity*. Dordrecht, The Netherlands: Springer.
- Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583–645.
- Fisher, C., Filby, N., Marliave, R., Cohen, L., Dishaw, M., Moore, J., & Berliner, D. (1978). *Teaching behaviors: Academic Learning Time and student achievement: Final report of Phase III-B, Beginning Teacher Evaluation Study*. San Francisco: Far West Laboratory for Educational Research and Development.
- Frick, T. (1990). Analysis of patterns in time (APT): A method of recording and quantifying temporal relations in education. *American Educational Research Journal*, 27(1), 180-204.
- Frick, T. (1997). Artificial tutoring systems: What computers can and can't know. *Journal of Educational Computing Research*, 16(2), 107-124.
- Frick, T., Chadha, R., Watson, C. & Zlatkowska, E. (2010a). College student perceptions of teaching and learning quality. *Educational Technology Research and Development*, 58(2), 115-136.

- Frick, T., Chadha, R., Watson, C., & Wang, Y. (2010b). Theory-based evaluation of instruction: Implications for Improving Student Learning Achievement in Postsecondary Education. In M. Orey, S. Jones, & R. Branch (Eds.), *Educational Media and Technology Yearbook*, 57-77.
- Frick, T., Chadha, R., Watson, C., Wang, Y. & Green, P. (2009). College student perceptions of teaching and learning quality. *Educational Technology Research and Development*, 57(5), 705-720.
- Greenspan, S. with Benderly, B. (1997). *The growth of the mind and the endangered origins of intelligence*. Reading, MA: Addison-Wesley.
- Keller, J. M. (1987). The systematic process of motivational design. *Performance & Instruction*, 26(9), 1-8.
- Kirkpatrick, D. (1994). *Evaluating training programs: The four levels*. San Francisco, CA: Berrett-Koehler.
- Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into Practice*, 41(4), 212-218.
- Kuh, G., Kinzie, J., Buckley, J., Bridges, B., & Hayek, J. (2007). Piecing together the student success puzzle: Research, propositions, and recommendations. *ASHE Higher Education Report*, 32(5). San Francisco: Jossey-Bass.
- Kulik, J. A. (2001). Student ratings: Validity, utility and controversy. *New Directions for Institutional Research*, 109, 9-25.
- Maccia, G. S. (1987). Genetic epistemology of intelligent natural systems. *Systems Research*, 4(1), 213-281.
- Merrill, M. D. (2002). First Principles of Instruction. *Education Technology Research & Development*, 50(3), 43-59.
- Merrill, M. D. (2007). A task-centered instructional strategy. *Journal of Research on Technology in Education*, 40(1), 33-50.
- Merrill, M. D. (2008). What makes e³ (effective, efficient, engaging) instruction? Keynote address at the AECT Research Symposium, Bloomington, IN.
- Merrill, M. D., Barclay, M., & van Schaak, A. (2008). Prescriptive principles for instructional design. In J. M. Spector, M. D. Merrill, J. van Merriënboer & M. F. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed.). New York: Lawrence Erlbaum Associates, 173-184.
- Montessori, M. (1965). *Spontaneous activity in education* (translated by F. Simmonds). New York: Schocken Books.
- Rangel, E., & Berliner, D. (2007). Essential information for education policy: Time to learn. *Research Points: American Educational Research Association*, 5(2), 1-4.
- van Merriënboer, J. J. G. & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.