International Journal of Engineering Education Vol. 27, No. 1, pp. 167–177, 2011
Printed in Great Britain

0949-149X/91 $3.00+0.00

A New Methodology for Assisting the Development of Instructional Awareness in Teaching a Large Engineering Class with Academically Diverse Students*

NING FANG
Department of Engineering and Technology Education, College of Engineering, Utah State University, Logan, UT 84322, USA.
E-mail: ning.fang@usu.edu

Instructional awareness is an instructor's clear and comprehensive understanding of his/her teaching pedagogy and practice. It affects teaching effectiveness and ultimately student learning. This paper presents a statistics-based, quantitative, and objective methodology to help develop instructional awareness in teaching a large engineering class with academically diverse students. The new methodology consists of three steps: collect data on student classwork and exam performance, identify student academic performance groups, and conduct paired-sample t-tests on each student group. A case study, which involved 236 undergraduate engineering students in three semesters, is provided to demonstrate how the methodology can be employed step by step. The results show that the methodology is valid to reveal whether or not an instructor's instructional strategies are more beneficial to one student group than another.

Keywords: instructional awareness; student academic performance groups; paired-sample t-tests; instructional strategies

1. Introduction

Instructional awareness is generally defined as an instructor's clear and comprehensive understanding of many aspects of his/her teaching pedagogy and practice. For example, how s/he teaches in the classroom, what technical content s/he teaches, and whether or not his/her instructional strategies help improve student learning [1–4]. In occasional cases, instructional awareness is referred to as "teaching awareness" [5]. As an increasing number of educational researchers and practitioners recognize the effect of "teachers make a difference" on student learning [6–8], growing attention has been paid to the study of instructional awareness and how it affects teaching effectiveness and ultimately student learning.

Weimer [9] presented a "five steps to better teaching" model:

(1) develop instructional awareness;
(2) seek inputs from colleagues and students;
(3) make choices about what ought to be changed and how to change it;
(4) implement the alterations;
(5) seek input about the alterations.

In that piece of work [9], Weimer provided a checklist for helping an instructor develop instructional awareness on how the instructor teaches in the classroom. The checklist includes seven questions, and each question contains several sub-questions. The seven questions are:

(1) What do you do with your hands?
(2) Where do you stand or sit?
(3) When do you move to a different location?
(4) Where do you move?
(5) Where do your eyes most often focus?
(6) What do you do when you finish one content segment and are ready to move on to the next?
(7) When do you speak louder/softer?

By observing or video recording an instructor's classroom lectures, these questions can be easily answered. However, these questions all focus on an instructor's behavior in the classroom; this is just only one of the many important aspects of instructional awareness.

1.1 Diverse student academic performance in large engineering classes

Felder and Brent [10] pointed out that students have "different levels of motivation, different attitudes about teaching and learning, and different responses to specific classroom environments and instructional practices." Moreover, there exists a semester-to-semester variation in student body and classroom composition. Even in the same semester, students can be significantly different from each other in terms of their academic background and performance [11]. Student difference is particularly significant in large, high-enrollment engineering classes, such as introductory foundational classes (e.g. Engineering Statics, Engineering Dynamics, and the Strength of Materials) that many freshman and sophomore students are required to take. Due

* Accepted 10 October 2010.

to significant differences of student academic background and performance, the final grades that students earned in these large courses can cover the full grade range from A, B, C, D, to Fail. It is a difficult task to teach these large classes because of the limited time that an instructor can spend on each student either inside or outside the classroom.

To maximize student learning outcomes in large classes, a variety of instructional strategies have been developed based on various learning theories and modern instructional technologies, for example, active learning [12, 13], cooperative learning [14], project-based and problem-based learning [15], multi-media instruction [16, 17], and technology-assisted learning [18, 19], to name a few. Usually, there is an overlap within these instructional strategies. For instance, one learning activity can be classified either as active learning or as project-based learning. Most often, an instructor should not rely on a single strategy (no one shoe fits all), and should adopt multiple instructional strategies to meet the needs of diverse students. Therefore, student learning outcomes largely depend on the combination of multiple instructional strategies.

A question arises regarding instructional awareness: How does an instructor know whether or not his/her instructional strategies work for diverse students in a large engineering class? In other words, do they work better for one group of students more than another? It is important to answer this question in order to maximize learning outcomes for all students in the class. If an instructor's instructional strategies are more beneficial to one group of students than another, the instructor should consider either redesigning or adjusting his/her instructional strategies, or taking additional measures to enhance the learning of those students in another group. Examples of additional measures may include the following:

(1) provide one-on-one, pre- and post-class tutoring for students;
(2) assign more representative technical problems;
(3) create additional course materials that include active and cooperative learning activities.

The answer to the above question can be obtained by interviewing or surveying students in the class. However, solely based on students' opinions and comments, the answer would be qualitative and subjective, rather than quantitative and objective. Two students in the same academic performance group could provide opposite opinions. A quantitative and objective answer will assist the instructor in developing a truer and more accurate understanding of his/her teaching pedagogy and practice.

1.2 Objective and research questions of the present study

The objective of this study is to establish a statistics-based, quantitative, and objective methodology for assisting the development of instructional awareness in teaching a large engineering class with academically diverse students. The three research questions of the present study include:

(1) Research Question 1: How does an instructor know whether or not his/her instructional strategies are more beneficial to one student group than another in a large engineering class?
(2) Research Question 2: If a quantitative methodology can be developed to answer the above question, is that methodology reliable and applicable in different semesters?
(3) Research Question 3: Can the developed methodology be cross-validated using another approach (method) that generates the same or similar findings?

1.3 Novelty of the present study

A variety of commonly-used literature databases were searched in this study that used various keywords, such as instructional awareness, teaching awareness, teacher self-awareness, teaching effectiveness, teaching assessment, and teacher evaluation. Such databases include the Education Resources Information Center, Science Citation Index, Social Science Citation Index, Engineering Citation Index, Academic Search Premier, the ASEE annual conference proceedings (1995–2010), and the ASEE/IEEE Frontier in Education conference proceedings (1995–2009).

The results show that the majority of relevant papers [such as 1–5, 9, 20] deal with general guidelines, check lists, or recommendations on what items included in the assessment and evaluation of an instructor's teaching effectiveness, or in a broader term "teacher evaluation." For example, in their paper on scholarly teaching, Fortenberry and Baber [20] presented a set of guidelines and checklists for peer classroom observations, such as an instructor's preparation and organization, content and knowledge, interaction with students, rapport and sensitivity, presentation (enthusiasm and clarity), and assessment of student learning, performance, and achievement. They emphasized what reflects self-assessment of instructors (i.e. instructional awareness) is an essential component of a systemic evaluation system. However, their paper provided no specific procedure on how to conduct reflective self-assessment.

There are studies focusing on the examination of teaching effectiveness on student learning. For example, Wright, Horn and Sanders [21] employed statistical mixed-model methodologies to conduct multivariate longitudinal analyses of a teacher's effect on student achievement. They found that differences in teacher effectiveness were a dominant factor affecting student academic gain. They suggested as a major component that teacher evaluation processes include a reliable and valid measure of the teacher effect on student academic growth over time.

From extensive literature search, however, no paper was found on developing a quantitative methodology or procedure to assist the development of instructional awareness in teaching large classes with diverse students. In particular, there is no existing study answering the question regarding how an instructor can distinguish whether or not his/her instructional strategies are more beneficial to one student group than another.

1.4 Logic structure and contents of this paper

This study presents:

(1) A detailed step-by-step description of the new methodology is provided. A large sophomore-year engineering class, Engineering Dynamics, is chosen as a case study that included the use of five diverse instructional strategies.
(2) How steps were done is illustrated, including how data were collected, how diverse student groups were identified, how statistical analysis was performed using the collected data, and what conclusions were drawn.
(3) After the methodology has been proven reliable (applicable) for multiple semesters, a different approach (method) is employed to cross-validate the methodology.
(4) Discussions are held on adopting the methodology by other instructors at different institutions.
(5) Two primary limitations of the methodology are also discussed.
(6) The research findings are summarized.

2. Methodology for assisting development of instructional awareness

2.1 Step-by-step description of methodology

The methodology developed in the present study is described in the following three steps:

Step 1: Collect data on each student's classwork performance and exam performance. In the present study, classwork performance is defined as the academic performance of a student participating in learning activities in regular classroom lectures throughout a semester. It can be quantified by the number (or percentage) of correct answers that a student makes in response to the instructor's technical questions in regular classroom lectures. Exam performance is defined as the academic performance of a student on mid-term and final exams. It is quantified by the scores that the student earned in these exams.

Step 2: Identify different academic performance groups (hereafter referred to as "student groups" or "groups") based on whether or not the performance of an individual student is above or below the average of all students in the class. For example, a student can belong to one of the following four groups:

Group I: classwork performance > the average classwork performance of all students, exam performance > the average exam performance of all students.

Group II: classwork performance < the average classwork performance of all students, exam performance > the average exam performance of all students.

Group III: classwork performance > the average classwork performance of all students, exam performance < the average exam performance of all students.

Group IV: classwork performance < the average classwork performance of all students, exam performance < the average exam performance of all students.

Group I consists of high-performing students, Groups II and III mid-performing students, and Group IV low-performing (or academically challenged, struggling) students.

Step 3: Conduct paired-sample t-tests on each student group to determine t-values (and their "significance" values) and the effect of size on each group. Based on the "significance" values and the magnitude of the effect size, determine whether or not an instructor's instructional strategies benefit one student group more than another.

2.2 Case study of applying the methodology in an Engineering Dynamics course

The above-described three-step methodology was applied in an Engineering Dynamics course the author has been teaching in the College of Engineering at Utah State University. Each year nearly 200 undergraduate students from four different departments take this required engineering course. The course covers a wide variety of foundational engineering concepts and principles; for example, motion, force and acceleration, work and energy, impulse and momentum, and vibrations of a particle and of a rigid body [22, 23]. The course is also an essential basis and fundamental building block for advanced studies in many subsequent courses, such as vibration, structural mechanics, system dynamics and control, and machine and structural designs.

To improve student learning of Engineering Dynamics, the author has adopted or developed diverse instructional strategies based on the active learning theory and its variety of best practices [24–26]. Five representative instructional strategies implemented in the course are described in the following paragraphs.

2.3 Instructional strategy #1

A Classroom Response System (colloquially called "clickers") was employed in each classroom lecture to improve active learning and student-instructor interactions and to provide the instructor with an immediate formative assessment of student learning [27]. Classroom Response System is a radio-frequency, a two-way communication system that comprises a set of hardware and computer software. The hardware consists of:

(1) a base that is connected to the instructor's computer in a classroom;
(2) a hand-held, portable clicker device, a wireless transmitter, typically $6'' \times 2'' \times 0.5''$.

Each clicker has a unique serial number that can be set up in advance to associate with the identification number of its owner (i.e., a student). When triggered, the base receives the real-time signals that each student submits from their transmitters.

During a lecture, the instructor poses a multiple-choice or true/false "clicker" question on a projection screen in front of the class, and then asks each student to push one of the five lettered buttons (A, B, C, D, and E) on the student's wireless clicker to anonymously respond to the question. The collective response from all students, the number or the percentage of the students who choose A, B, C, D, or E, is immediately displayed on the screen, so both the instructor and students can perceive the collected response. In other words, clickers provide both the instructor and students with immediate feedback and real-time assessment of learning; thus teaching can also be adjusted in real-time.

The computer software associated with the clicker system automatically generates two reports for each lecture:

(1) lecture summary report listing each clicker question, number of student responses, maximum score and average score for each question;
(2) student voting report listing response of each student to each clicker question.

Therefore, after the class, the instructor can use these reports to analyze in detail the academic performance of each student.

2.4 Instructional strategy #2

A set of interactive computer visualization modules that covered the most important Dynamics concepts and principles in each chapter of the textbook [22] were developed and provided to students as supplementary learning materials. For example, one module called "fire the mortar and hit the target" focused on the analysis of velocity and acceleration in a projectile motion. By changing the initial angle and initial velocity of a projectile, the students can see how the velocity varies during a projectile motion. Students can run those computer visualization modules at their own pace and at anytime (24/7).

2.5 Instructional strategy #3

After each classroom lecture, students were required to write a short paragraph (called keynotes) describing what they had learned in the lecture. Student-created keynotes were frequently reviewed during classroom lectures, so students would not get lost from numerous learning topics in the course.

2.6 Instructional strategy #4

An electronic Classroom Management System (Blackboard™) was employed to facilitate student-student and student-instructor communications after class. In addition to being a central place for the instructor to collect and grade all student homework assignments, the system enabled students to exchange emails with each other and with the instructor. Many students used Blackboard™ to discuss with the instructor technical problems when the students cannot meet the instructor in person in the instructor's office hour.

2.7 Instructional strategy #5

After-class Help Sessions were held daily to answer students' questions and help them solve technical problems. These help sessions were primarily conducted by the teaching assistants on the course.

Clearly, the final outcomes of student learning were affected by the combination of all five instructional strategies. A question regarding instructional awareness arises: How does the combined use of these strategies affects diverse student academic performance groups? Sections 3–6 provide a detailed step-be-step description of how this question is answered.

To avoid potential conflicts and interference between teaching and research, no efforts were made to study the effect of a single instructional strategy on student learning. To study the individual effect, only one single instructional strategy would be allowed in teaching in one semester. This would reduce student learning outcomes (as compared to

Table 1. Student demographics

	Major *			Sex	
	MAE	**CEE**	**Other**	**Male**	**Female**
Semester #1 (n = 128)	72 (56.3%)	34 (26.5%)	22 (17.2%)	108 (84.4%)	20 (15.6%)
Semester #2 (n = 58)	22 (37.9%)	20 (34.5%)	16 (27.6%)	51 (87.9%)	7 (12.1%)
Semester #3 (n = 50)	23 (46.0%)	19 (38.0%)	8 (16.0%)	42 (84.0%)	8 (16.0%)
Total (n = 236)	117 (49.6%)	73 (30.9%)	46 (19.5%)	201 (85.2%)	35 (14.8%)

* MAE:Mechanical and aerospace engineering.
CEE: Civil and environmental engineering.
Other: Biological engineering, general engineering, pre-engineering, undeclared majors, etc.

using multiple instructional strategies) and unfair to students in that particular semester.

3. Step 1—collecting data

As stated in Instructional Strategy #1 in the previous section, the clickers provide a fast and convenient tool for an instructor to collect data on academic performance of each individual student during regular classroom lectures. The data can be analyzed after class. Therefore, a student's classwork performance was quantified by the average "clicker" score that the student earned in regular classroom lectures. In the present study, the score that a student earned after-class homework assignments was not taken into account because the after-class homework score might not reflect the true learning gain of an individual student. For instance, students could work in groups or consult a homework solution manual.

A student's exam performance in the Engineering Dynamics course was quantified by the average exam score that the student earned in four exams: three mid-term exams and one final exam. Each mid-term exam covered two chapters of the textbook [22]. The final exam was comprehensive and covered nine chapters of the textbook [22].

Data on student classwork and exam performance were collected from a total of 236 students in three semesters: 128 students in Semester #1, 58 students in Semester #2, and 50 students in Semester #3. Table 1 shows student demographics.

Table 1 shows that most of the 236 students were either from the mechanical and aerospace engineering major (49.6%) or from the civil and environmental engineering major (30.9%). The vast majority of students were male (85.2%), and the female students accounted for 14.8%.

4. Step 2—identifying student groups

Based on their classwork and exam performance, students were then divided into four groups: I, II, III, and IV. The way in which these four student groups were formed has been described in Section 2.1.

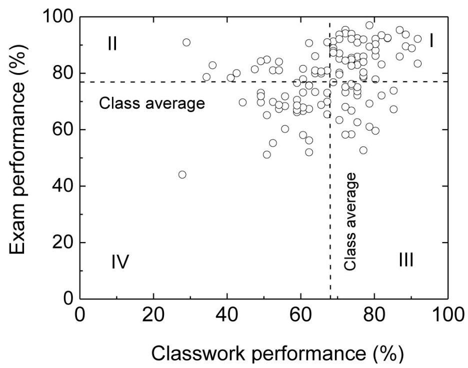

Fig. 1. Student groups in Semester #1 (n = 128).

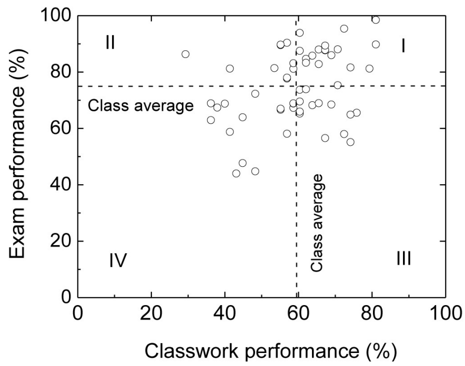

Fig. 2. Student groups in Semester #2 (n = 58).

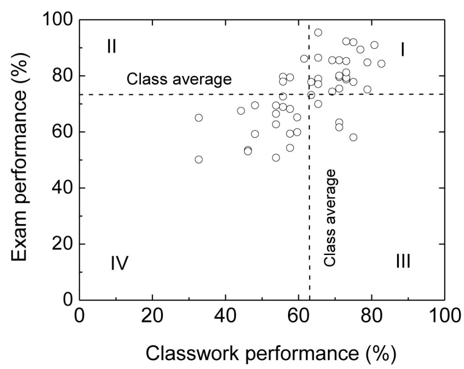

Fig. 3. Student groups in Semester #3 (n = 50).

Figures 1, 2 and 3 show student groups in Semesters #1, #2, and #3, respectively. Tables 2–4 further provide descriptive statistics for each student group

Table 2. Descriptive statistics for each student group (Semester #1)

Student Group	Performance	Mean	Standard Deviation	Standard Error Mean
I	Clicker score	77.25	6.36	0.91
	Exam score	88.12	4.88	0.70
II	Clicker score	56.45	11.16	2.50
	Exam score	83.06	3.81	0.85
III	Clicker score	75.84	4.46	0.93
	Exam score	68.94	7.33	1.53
IV	Clicker score	56.06	8.83	1.47
	Exam score	67.57	8.03	1.34

Table 3. Descriptive statistics for each student group (Semester #2)

Student Group	Performance	Mean	Standard Deviation	Standard Error Mean
I	Clicker score	67.33	8.34	1.87
	Exam score	84.87	10.19	2.28
II	Clicker score	52.24	9.48	3.00
	Exam score	83.85	4.82	1.52
III	Clicker score	66.58	5.96	1.65
	Exam score	65.68	5.96	1.65
IV	Clicker score	47.95	8.63	2.23
	Exam score	63.21	8.55	2.21

Table 4. Descriptive statistics for each student group (Semester #3)

Student Group	Performance	Mean	Standard Deviation	Standard Error Mean
I	Clicker score	72.24	5.21	1.09
	Exam score	82.90	6.09	1.27
II	Clicker score	57.69	2.72	1.36
	Exam score	80.77	3.61	1.81
III	Clicker score	69.23	4.71	2.11
	Exam score	65.23	6.20	2.77
IV	Clicker score	50.96	8.19	1.93
	Exam score	61.97	7.18	1.69

Table 5. Number (percentage) of students in each group

Student Group	Semester #1 (n = 128)	Semester #2 (n = 58)	Semester #3 (n = 50)
I	n_1 = 49 (38.3%)	n_1 = 20 (34.5%)	n_1 = 23 (46.0%)
II	n_2 = 20 (15.6%)	n_2 = 10 (17.2%)	n_2 = 4 (8.0%)
III	n_3 = 23 (18.0%)	n_3 = 13 (22.4%)	n_3 = 5 (10.0%)
IV	n_4 = 36 (28.1%)	n_4 = 15 (25.9%)	n_4 = 18 (36.0%)

for three semesters. Table 5 shows the number (percentage) of students in each group. A consistent pattern of student distribution in each group among all three semesters can be found in Table 5. This consistent pattern is: Group I has the largest percentage (34.5%–46.0%) of students; Group IV the second largest (25.9%–36.0%); then Group III (10.0%–22.0%); and finally Group II the smallest (8.0%–17.2%). Because the data was collected from three different semesters, this consistent pattern is not coincident but represents the characteristics of student composition in the College of Engineering at the author's university.

5. Step 3—paired-sample t-tests on each student group

As is well known, t-tests are employed to determine whether there is a statistically significant difference between two sample means [28, 29]. There are two different types of t-tests: independent means t-tests (also called independent samples t-tests) and dependant means t-tests (also called paired-samples t-tests). The first t-test can be used when there are two experimental conditions each of which has different participants. The second can be used when there are two experimental conditions each of which

Table 6. Paired-samples t-tests for each student group (Semester #1)

	Paired Differences			t-test		
Pair in Student Group	Mean	Standard Deviation	Standard Error Mean	t-value	Degree of freedom (df)	Significance (2-tailed)
I	–10.87	7.26	1.03	–10.48	48	0.000
II	–26.61	12.63	2.82	–9.42	19	0.000
III	6.90	8.85	1.85	3.74	22	0.001
IV	–11.51	10.94	1.82	–6.32	35	0.000

Table 7. Paired-samples t-tests for each student group (Semester #2)

	Paired Differences			t-test		
Pair in Student Group	Mean	Standard Deviation	Standard Error Mean	t-value	Degree of freedom (df)	Significance (2-tailed)
I	–17.55	8.33	1.86	–9.42	19	0.000
II	–31.61	11.09	3.51	–9.01	9	0.000
III	0.89	10.64	2.95	0.30	12	0.767
IV	–15.26	11.30	2.92	–5.23	14	0.000

Table 8. Paired-samples t-tests for each student group (Semester #3)

	Paired Differences			t-test		
Pair in Student Group	Mean	Standard Deviation	Standard Error Mean	t-value	Degree of freedom (df)	Significance (2-tailed)
I	–10.66	7.31	1.52	–7.00	22	0.000
II	–23.07	1.35	0.68	–34.15	3	0.000
III	4.01	10.89	4.87	0.82	4	0.457
IV	–11.01	9.36	2.21	–4.99	17	0.000

has the same participants. In the present study, the same group of students in a semester participated in both regular classroom lectures and exams. Therefore, paired-samples t-tests are employed.

The t-value is calculated as [30]

$$t = \frac{\bar{D} - \mu_D}{SD/\sqrt{N}} \tag{1}$$

where $\bar{D}$ is the observed difference between sample means, μ_D is the expected difference between population means (if null hypothesis is true), SD is standard deviation, and N is the number of samples. The term $SD/\sqrt{N}$ represents the standard error of the differences between two sample means. If the null hypothesis is true, then there is no difference between the population means (i.e. $\mu_D = 0$).

Tables 6–8 show the results of paired-samples t-tests for each student group in three semesters. The negative t-values indicate that the mean value of clicker scores is less than the mean value of exam scores, or student academic performance improved in exams. From the "significance" values reported in the last columns in Tables 6–8, there exist statistically significant differences ($p < 0.001$) between classwork performance mean and exam performance mean for nearly all student groups in all three semesters (with the only two exceptions for Student Group III in Semesters #2 and #3).

The magnitude of effect size was further determined to provide an objective measure of the importance of an effect. The effect size (r-value) was calculated as [31]

$$r = \sqrt{\frac{t^2}{t^2 + df}} \tag{2}$$

The widely accepted suggestions about the magnitude of effect size are as follows [32]:

- r = 0.10 (small effect): the effect accounts for 1% of the total variance;
- r = 0.30 (medium effect): the effect accounts for 9% of the total variance;
- r = 0.50 (large effect): the effect accounts for 25% of the total variance.

Table 9 shows the effect size for each student group in three semesters. As seen from Table 9, a large effect size exists for nearly all student groups in all three semesters (except for Student Group III in Semesters #2 and #3). This finding is consistent with the analysis based on the "significance" values.

More important, a comparison among the mag-

Table 9. Effect size (r-value) for each student group

Student Group	Semester #1	Semester #2	Semester #3
I	0.83	0.91	0.83
II	0.91	0.95	1.00
III	0.62	0.09	0.38
IV	0.73	0.81	0.77

nitudes of r-values shows a consistent pattern of the magnitude of effect size in three semesters: Group I > Group IV. Group II is not included in the comparison due to relatively small number of students in Group II (totaling 34 students in three semesters refer to Table 5). Group III included is not included in the comparison because of statistically insignificant differences between classwork performance mean and exam performance mean in Semesters #2 and #3.

The above findings imply that the author's instructional strategies had a more profound impact on student learning for Group I than for Group IV. Worth noting is that the classwork and exam performance of students in Group I are both above the class average; whilst the classwork and exam performance of students in Group IV were both below the class average. Therefore, it can be concluded that the author's instructional strategies were more beneficial to high-performing students (i.e. students in Group I) than low-performing, academically struggling students (i.e. students in Group IV). In other words, under the instruction of the author, all students in the class learned, with high-performing students having greater learning gain than low-performing students. But, there is room for improvement of low-performing students by offering more instructional interventions (in addition to those five instructional strategies) to adapt to those student's learning styles, cognitive and meta-cognition skills, time management skills, prior course work and achievements, and so on. On the other hand, students themselves must also take proactive learning and find out the best ways to learn, so teaching and learning can be mutually beneficial and enhanced.

6. Cross-validation of the methodology

The analysis described above reveals a consistent pattern of the effect in the author's instructional strategies on academically diverse student groups in three different semesters. This supports the reliability of the methodology, or the applicability of the methodology in different semesters. The next question is how to prove the validity of the methodology, or determine if another approach (method) can be used to generate the same or similar findings?

Let us study whether or not there is a statistically significant correlation between a student's Grade Point Average (GPA) and the t-value. Typically GPA is a comprehensive representation of student performance. A high GPA is generally associated with high performance of a student, and a low GPA comes with low performance. The t-value is an indication of student learning as the result of the diversified instructional strategies.

The t-value calculated in the traditional way (i.e. Equation 1) is for a group of participants. A mathematical treatment must be applied first to determine the t-value for an individual student (participant), so a sufficient number of data points can be generated to correlate GPA and the t-value in a meaningful statistical sense.

To determine the t-value for an individual participant, Equation (1) is rewritten as

$$t = \frac{\bar{D}}{SD/\sqrt{N}} = \frac{1}{N}\left(\frac{D_{S1} + D_{S2} + \dots + D_{SN}}{SD/\sqrt{N}}\right) = \frac{1}{N}\left(\frac{D_{S1}}{SD/\sqrt{N}} + \frac{D_{S2}}{SD/\sqrt{N}} + \dots + \frac{D_{SN}}{SD/\sqrt{N}}\right) \quad (3)$$

where D_{S1}, D_{S2}, ..., D_{SN} are the difference between the two samples in Sample Pairs 1, 2, .. n, respectively; and the terms

$$\frac{D_{S1}}{SD/\sqrt{N}}, \frac{D_{S2}}{SD/\sqrt{N}}, \dots, \frac{D_{SN}}{SD/\sqrt{N}}$$

represent the t-value for Sample Pair 1, 2, .. n, respectively. These latter terms can be used to calculate the t-value for each individual student in the group to which the student belongs.

The term $SD/\sqrt{N}$ represents the standard error of the differences between two sample means. Therefore, the t-value for an individual student (participant) can also be calculated as:

$$\text{The t-value for an individual student} = \frac{\text{AC} - \text{AE}}{\text{SE}} \quad (4)$$

where AC is the student's average clicker score, AE is the student's average exam score, and SE is the stand error of the differences between two sample means in the group the student belongs.

Figures 4–6 show each student's GPA vs. t-value in three semesters. One open circle represents one student in these figures. A negative t-value indicates that a student's average clicker score was less than his/her average exam score, or the student's academic performance improved in exams. For example, a student with the t-value of –20 performs better than another student with the t-value of –10.

As seen from Figures 4–6, particularly from Figures 5 and 6, a linear relationship exists between students' GPA and t-value. An increased GPA is associated with a decreased t-value, and vice versa. Table 10 further shows the Pearson correlation between students' GPA and t-value. A statistically significant correlation can be found in all three semesters: $p < 0.007$ for Semester #1, $p < 0.001$ for Semester #2, and $p < 0.010$ for Semester #3. This means the author's instructional strategies (indicated by the t-value) impact different students (represented by GPA). This finding is similar to what has been concluded by using the methodology described before. In other words, the methodology is cross validated.

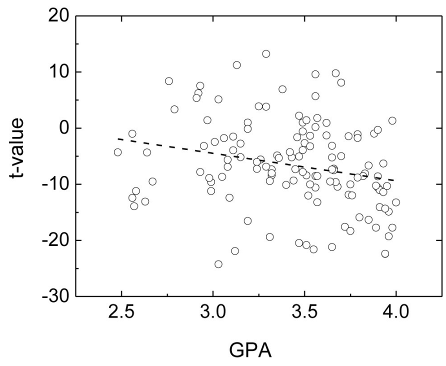

Fig. 4. Students' GPA vs. t-value in Semester #1 (n = 128).

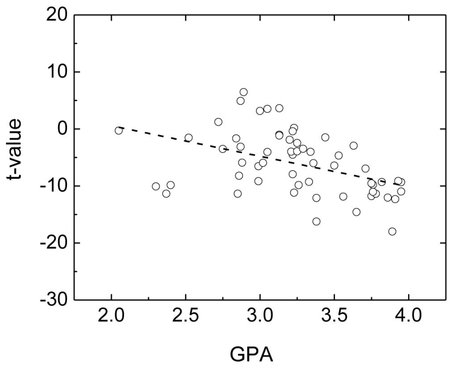

Fig. 5. Students' GPA vs. t-value in Semester #2 (n = 58).

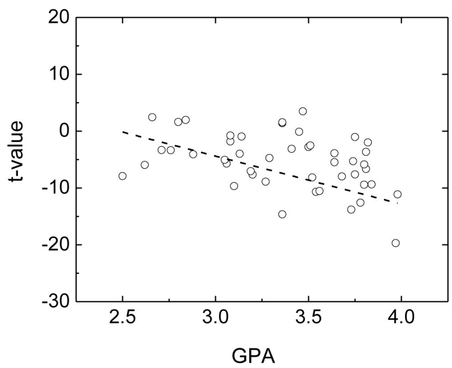

Fig. 6. Students' GPA vs. t-value in Semester #3 (n = 50).

Table 10. Correlation between students' GPA and t-value

Semester	Number of students	Pearson Correlation	Significance (2-tailed)
#1	128	–0.237*	0.007
#2	58	–0.441*	0.001
#3	50	–0.360**	0.010

* Correlation is significant at the 0.01 level (2-tailed).
** Correlation is significant at the 0.05 level (2-tailed).

7. Discussions and limitations

The methodology described in this paper was developed from the author's teaching practice at a large public university. However, the methodology can also be easily adopted by instructors at other institutions to help develop instructional awareness in teaching large engineering classes with diverse students. Simply follow the aforementioned three steps.

In the present study, clickers were used in Step 1 as a fast and convenient tool to collect data on student classwork performance. If no clickers are used at other institutions, an alternative method may be used to collect data on student classwork performance. For example, an instructor can deign his/her own technical problems for students to solve in or after classes. A care must be taken when using homework problems provided in the textbook to judge student classwork performance because the solution manual is easily accessible to students nowadays.

The primary limitation of the developed methodology is how it collects data in multiple semesters, which adds to the workload for an instructor. It is suggested that data be collected for at least three consecutive semesters to accumulate a sufficient number of data points for statistical analysis. In addition, the instructor's strategies should not be changed significantly throughout these semesters; otherwise, a consistent pattern of the impact of the instructor's strategies on diverse students might not be found from statistical analysis.

The second possible limitation of the developed methodology is no statistically reasonable conclusions can be made for mid-performing student groups. For example, in Step 3 in the present study, students in Group III could not be included in the comparison of the magnitude of effect size (r-values), because the difference between classwork performance mean and exam performance mean in Semesters #2 and #3 for students in Group III was found to be statistically insignificant. In this case, even if the data were collected and analysis was done, no statistically reasonable conclusions could be drawn. In fact, the performance of students in mid-performing groups (II and III) turns out to be unstable and inconsistent from semester to semester compared with the performance of students in top-performing and low-performing groups (I and IV).

8. Conclusions

Instructional awareness is an instructor's clear and comprehensive understanding of his/her teaching pedagogy and practice. It directly affects teaching effectiveness and ultimately student learning outcomes. In the present study, a statistics-based, quantitative, and objective methodology has been established to help develop instructional awareness in teaching a large engineering class with diverse students. An Engineering Dynamics course has been used as a case study to illustrate step by step how the new methodology was employed. The following paragraphs summarize the answers to the three research questions of the present study.

Research Question 1: How does an instructor know whether or not his/her instructional strategies are more beneficial to one student group than another in a large engineering class?

Answer: Step 1: collect data on each student's classwork performance and exam performance. Step 2: identify student academic performance groups based on the collected data. Step 3: conduct paired-sample t-tests on each student group. By analyzing t-values and effect sizes, an instructor can determine whether or not his/her instructional strategies are more beneficial to one student group than another.

Research Question 2: If a quantitative methodology can be developed to answer the above question, is that methodology reliable and applicable in different semesters?

Answer: Yes, it is. Data have been collected from a total of 236 students in three semesters: 128 students in Semester #1, 58 students in Semester #2, and 50 students in Semester #3. The results of paired-sample t-tests have shown that a large effect size exists for nearly all student groups in all three semesters. There exists a consistent pattern of the magnitude of effect size in all three semesters, i.e. Group I > Group IV. In other words, the author's instructional strategies had a more profound impact on student learning for high-performing students (Group I) than for low-performing students (Group IV).

Research Question 3: Can the developed methodology be cross-validated using another approach (method) that generates the same or similar findings?

Answer: Yes, it can be validated by investigating the correlation between a student's Grade Point Average (GPA) and t-value. Results show a statistically significant correlation between a student's GPA and t-value for all three semesters: $p < 0.007$ for Semester #1, $p < 0.001$ for Semester #2, and $p < 0.010$ for Semester #3. This means the author's instructional strategies (indicated by the t-value) impact different students (represented by GPA), which is a similar finding that has been made by the methodology.

The new methodology can be easily adopted by instructors at other institutions to help develop instructional awareness in teaching large classes with diverse students. However, it has two limitations. First, it requires data to be collected in multiple semesters, which adds workload for an instructor. Second, there is a possibility that no statistically meaningful conclusions can be made for mid-performing student groups.

References

1. M. Weimer, *Learner-Centered Teaching: Five Key Changes to Practice,* John Wiley & Sons, Inc., Hoboken, NJ, 2002.
2. K. G. Lewis, Gathering data for the improvement of teaching: what do I need and how do I get it? *New Directions for Teaching and Learning*, **1991**, 2006, pp. 65–82.
3. C. Finelli, M. Ott, A. C. Gottfried, C. Hershock, C. O'Neal and M. Kaplan, Utilizing instructional consultations to enhance the teaching performance of engineering faculty. *Journal of Engineering Education*, **97**, 2008, pp. 397–411.
4. R. S .J. Reyes, The efficiency and effectiveness of engineers as full-time and part-time teachers in selected colleges of engineering of higher learning: an assessment. *The 29th ASEE/IEEE Frontiers in Education Conference*, November 10–13, San Juan, Puerto Rico 1999.
5. E. I. Ko and J. R. Hayes, Teaching awareness of problem-solving skills to engineering freshmen. *Journal of Engineering Education*, **83**, 1994, pp. 1–5.
6. S. D. Tu, Total engineering education: The way to shape our future leadership. *Global Journal of Engineering Education*, **10**, 2006, pp. 117–128.
7. R. M. Felder, A. Rugarcia and J. E. Stice, The future of engineering education V. assessing teaching effectiveness and educational scholarship**.** *Chemical Engineering Education,* **34**, 2000, pp. 198–207.
8. G. K. Cook, A vector approach to the assessment of undergraduate engineering courses. *International Journal of Engineering Education*, **16**, 2000, pp. 3–11.
9. M. Weimer, *Improving College Teaching: Strategies for Developing Instructional Effectiveness,* John Wiley & Sons, Inc., Hoboken, NJ, 1990.
10. R. M. Felder and R. Brent, Understanding student difference. *Journal of Engineering Education*, **94**, 2005, pp. 57–72.
11. K. Herr, *Improving teaching and learning in large classes: a practical manual,* Office of Instructional Services, Colorado State University, Fort Collins, CO, 1985.
12. E. Graaff, G. N. Saunders-Smits and M. R. Nieweg, *Research and Practice of Active Learning in Engineering Education*, Palls Publications, Amsterdam, Holland, 2005.
13. M. Prince, Does active learning work? A review of the research. *Journal of Engineering Education*, **93**, 2004, pp. 223–231.
14. L. Morell, R. Buxeda, M. Orengo and A. Sanchez, After so much effort: Is faculty using cooperative learning in the classroom? *Journal of Engineering Education*, **90**, 2001, pp. 357–362.
15. K. A. Smith, S. D. Sheppard, D. W. Johnson and R. T. Johnson, Pedagogies of engagement: classroom-based practices. *Journal of Engineering Education*, **94**, 2005, pp. 87–101.
16. M. Budhu, Enhancing instruction using interactive multimedia simulations, *Simulation*, **76**, 2001, pp. 222–231.
17. L. Whitman, D. Malzahn, Van Madhavan, G. Weheba and

K. Krishnan. Virtual reality case study throughout the curriculum to address competency gaps. *International Journal of Engineering Education*, **20**, 2004, pp. 690–702.
18. L. A. Van Dijk, G. C. Van Den Berg and H. Van Keulen, Interactive lectures in engineering education. *European Journal of Engineering Education*, **26**, 2001, pp. 15–28.
19. S. Rothberg, F. Lamb and A. Wallace, Computer assisted learning in engineering degree Programmes: a survey at the end of the 20th century. *International Journal of Engineering Education*, **17**, 2001, pp. 502–511.
20. N. Fortenberry and T. Baber, Evaluating instructional scholarship in engineering. *Proceedings of the 2008 American Society for Engineering Education Annual Conference & Exposition*, Pittsburgh, PA, 2008.
21. S. P. Wright, S. P. Horn and W. L. Sanders, Teacher and classroom context effects on student achievement: implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, **11**, 1997, pp. 57–67.
22. R. C. Hibbeler, *Engineering Mechanics Dynamics* (12th edition), Pearson Prentice Hall, Upper Saddle River, NJ, 2009.
23. A. Bedford and W. Fowler, *Engineering Mechanics Dynamics* (5th edition), Prentice Hall, Upper Saddle River, NJ, 2009.
24. B. Self and R. Redfield, New approaches in teaching undergraduate dynamics. *Proceedings of the 2001 American Society for Engineering Education Annual Conference & Exposition*, Albuquerque, NM, 2001.
25. M. B. Rubin and E. Altus, An alternative method for teaching dynamics. *International Journal of Engineering Education*, **16**, 2000, pp. 447–456.
26. P. J. Cornwell. Teaching dynamics using modern tools. *Computers in Education*, **6**, 1996, pp. 18–24.
27. N. Fang, An electronic classroom response system for an engineering dynamics course: student satisfaction and learning outcomes. *International Journal of Engineering Education*, **25**, 2009, pp. 1059–1067.
28. D. B. Wright, *First Steps in Statistics*, Sage Publications Ltd., London (2002).
29. B. H. Cohen, *Explaining Psychological Statistics* (2nd edition), John Wiley & Sons, Inc., New York, NY, 2000.
30. A. Field, *Discovering Statistics Using SPSS* (2nd edition), Sage Publications Ltd., London, 2005.
31. R. L. Rosnow and R. Rosenthal, *Beginning Behavioural Research: A Conceptual Primer* (5th edition), Preason/Prentice Hall, Englewood Cliffs, NJ (2005).
32. J. Cohen, A power primer. *Psychological Bulletin*, **112**, 1992, pp. 155–159.

Ning Fang is an Associate Professor in the College of Engineering at Utah State University, USA. He teaches Engineering Dynamics. His areas of interest include computer-assisted instructional technology, curricular reform in engineering education, the modeling and optimization of manufacturing processes, and lean product design. He earned his Ph.D., M.S., and BS degrees in mechanical engineering and is the author of more than 60 technical papers published in refereed international journals and conference proceedings. He is a Senior Member of the Society for Manufacturing Engineering and a member of the American Society of Mechanical Engineers. He is also a member of the American Society for Engineering Education and a member of the American Educational Research Association.