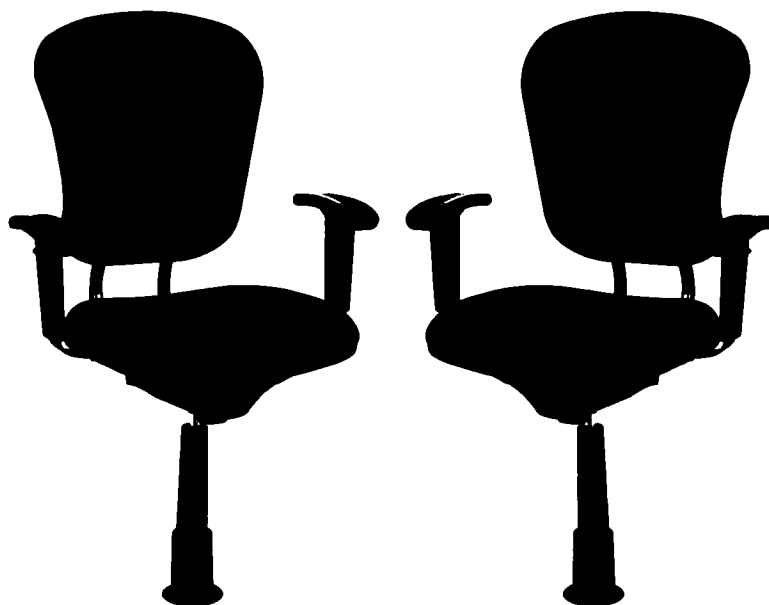


Swivel Seating in Large Lecture Theaters and Its Impact on Student Discussions and Learning

By Craig A. Ogilvie



This article describes a remodeled, large lecture theater that has swivel seats to facilitate face-to-face peer discussions, and the impact this has on student learning.

Well-designed university buildings and physical environments have a documented positive impact on student participation, engagement, and feelings of support and belonging (Strange and Banning 2001). These factors are known to improve learning; however, it is hard to document the direct impact architecture has on student learning outcomes (Strange and Banning 2001; Dane 2004). Information on how specific design choices impact student learning would be tremendously beneficial to the development of new projects, and outcome data from prior remodeling projects could be used to inform the myriad choices architects and universities must make for new designs.

This paper compares two different designs of remodeled, large lecture rooms and the impact these designs had on student learning. Funds were available to renovate the two large lecture theaters used by the physics and astronomy department at Iowa State University. As a department, we were

not able to decide between two different lecture-theater designs: one with traditional tiered rows and one with swivel seating to facilitate face-to-face discussions during lectures (described in more detail below). As a compromise, both designs were built: one theater with traditional seating and one with swivel seating. Hence, I had the opportunity to compare the impact that discussion seating had on student learning.

There is a tremendous amount of information available about the design of learning spaces—both formal classrooms and more informal gathering spaces at a university (Strange and Banning 2001; Oblinger 2005). Design guidelines are occasionally complemented by postconstruction quantitative data on enrollment growth, retention, increase in morale, ease of recruiting faculty, increase in grants, and participation of undergraduates in science research (Project Kaleidoscope 1998). While a great deal of focus has been on improving the visual/audio aspects of large lecture theaters, the fundamental layout of these rooms has not

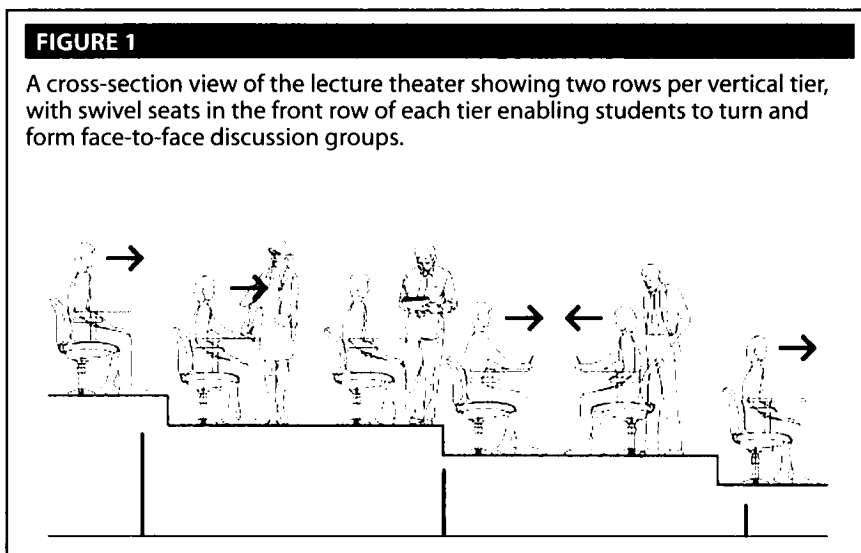
changed dramatically over the years. The traditional design consists of rows of seats, each row at a slightly different elevation (tiered rows) with the professor lecturing from the front. The focus is on the professor, with students' role of taking notes cued by the familiar seating arrangement (Jamieson, Dane, and Lippman 2005). The traditional large lecture theater is essentially designed for one-way delivery of information. Implementing pedagogical advances, such as in-lecture small-group discussions, is possible but traditional design does not necessarily encourage or support these techniques.

Within the constraints of traditional tiered rows in large lecture theaters, considerable work has been done on optimizing the line of sight for students and improving the acoustics, overall lighting, temperature, aesthetics, size of seats, desktop arrangements, material finishes, and the use of textures and colors (Lang 1996). Documentation is available as "golden rules" (Bartlett 2003; Bartlett and Thomason 1983) and design standards (UMBC 2000). These guidelines have been informed by surveys of students and professors (Fleming and Starr 1999) that link preferences to specific design elements in the classroom (Douglas and Gifford 2001).

Improved pedagogy in large lecture theaters

Students learn more if they can discuss new ideas with their peers—explaining, justifying, and arguing to support their thinking (Johnson, Johnson, and Smith 1991; Bransford, Brown, and Cocking 2000). As an example in physics, a question might ask students to compare the acceleration of a thrown ball when it is at the top of its arc compared to when the ball is about to hit the ground. Research has shown that many students have a misconception that the acceleration is zero at the top, and that this misconception persists even after attending lectures and solving numerical problems on the topic (Halloun and Hestenes 1985).

Peer discussions within large lectures (Mazur 1997) help reduce these misconceptions. In this pedagogy, the professor lectures on a topic for 10–15



minutes, then asks a cognitive conflict question with multiple-choice answers, including distracters chosen to represent common misconceptions. Each student is accountable for an answer, typically via a handheld "clicker" (Mazur 1997), where after working through the question individually, each student commits to an answer by punching in a multiple-choice option on the clicker. After this first round of answers, students turn to their neighbors to discuss the question with a small group of peers, before recommitting to an answer by clicking again.

Swivel seating

How can this technique be optimized? Does the design of the seating arrangement affect the learning gained during peer discussions? Using a renovation project to our large lecture theaters, I strove to facilitate face-to-face discus-

sions by having two rows of seats for every vertical elevation, or tier. The front row of seats in each tier could swivel, so that three to four students could easily form groups for discussions. This swivel-seat design for small groups within large lectures has been suggested in the literature (Jamieson, Dane, and Lippman 2005), but as far as I am aware, this seating arrangement has not been implemented elsewhere in large lecture theaters.

Figure 1 shows a cross-section view of the swivel-seat lecture theater. Students face forward during the lecture, then use the swivel seats to have face-to-face discussions with their peers. The seats can swivel only 315° to prevent students from spinning. A photograph of the completed lecture theater is shown in Figure 2. The capacity of this room is 117 students. A

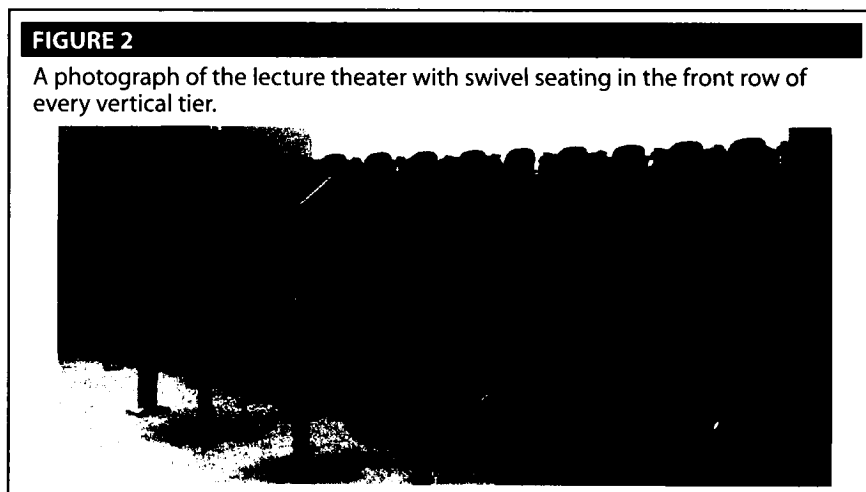


FIGURE 3

A photograph of the lecture theater with traditional tiered seating, i.e., one row for every vertical tier.



second lecture theater was renovated at the same time. The design for this theater was traditional, with one row of seats for every vertical tier as shown in Figure 3. The capacity of this room is 270 students. A plan view of both rooms is shown in Figure 4. Both rooms were designed to have good line of sight to two large projection screens placed above a traditional blackboard. Both rooms have good acoustics and lighting and a light, relaxed color scheme.

Research design

The data presented in this study come from the spring 2006 semester of a sophomore, calculus-based physics course that I taught. Three hundred and fifty students took the course. The course met for three lectures each week,

one recitation, and one lab. The recitations used a mixture of physics tutorials (McDermott and Shaffer 1992; Shaffer and McDermott 1992) and context-rich multifaceted problem solving (Heller, Keith, and Anderson 1992; Heller and Hollabaugh 1992) designed to increase problem-solving skills.

Each lecture was presented three times, back-to-back. The first lecture was held in the traditional lecture theater, the next two were held in the lecture theater with swivel seats. The lecture notes were PowerPoint slides, so identical content was delivered in each room.

The active-learning format of the lecture was described in the previous section: approximately 10 minutes of mini-lecture about an idea, followed by

a conceptual question (referred to as a *concepTest* [Mazur 1997] for the rest of this paper; for a sample *concepTest* that was asked on heat transfer and temperature changes, see Figure 5).

There were always two rounds for each question; during the first round individuals work on the question, then commit to an answer via an infrared clicker. The first round typically lasted 60 to 80 seconds, depending on the pace at which answers came in. At the start of the second round, students formed small groups amongst their neighbors to discuss the question, debate explanations, and come to some consensus. Each student then recommitted to an answer via their clicker. The second round typically lasted 90 to 120 seconds, including the small group discussions.

At the end of the first round, the frequency distribution of answers was shown to students. If no answer dominated, then the subsequent discussion during the second round was intense. The question was now even more of a puzzle—which answer was correct and why? The likelihood of having different viewpoints within each group was high and students had received the implicit support that others in the class chose the same answer that they did. If one answer had a large majority after the first round, then the discussion was shorter, less intense, with perhaps some peer explanations to students who may have chosen one of the less common answers.

I strove for a mix of questions; some that would produce no clear majority after the first round, complemented by some easier questions with a likely dominant answer. The latter were included because of a documented negative effect on students' perceived common sense if they are asked repeated "cognitive conflict" questions in which they struggle to give the correct answer (Redish, Saul, and Steinberg 1998). The answer to the *concepTest* was revealed after the second round, often producing groans, or exclamations of "yes!" from students. The second-round answers provided students with a small amount of extra credit (0.3% of final grade).

FIGURE 4

A plan view of the two renovated lecture theaters. The theater on the right has two rows for every vertical tier, with swivel seating in the front row of each tier. The theater on the left is a more traditional layout with one row for each vertical tier.

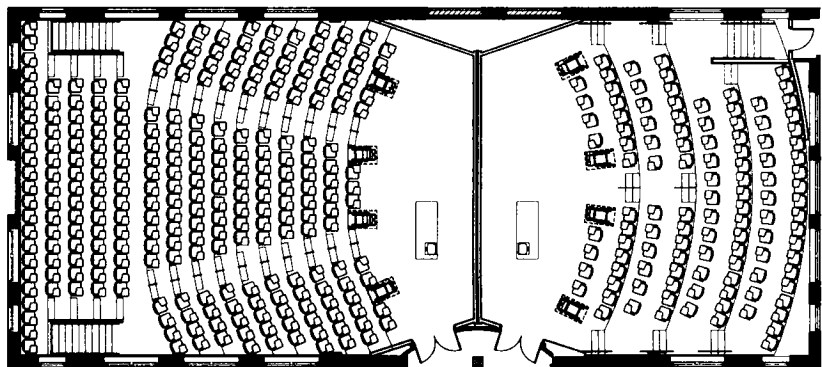


FIGURE 5

Sample conceptTest.

- You have some gold and the same mass of plastic
 - Plastic has twice the specific heat capacity as the gold
 - You transfer the same amount of heat to both gold plastic
- How do the temperature changes compare?

(1) $\Delta T_{\text{plastic}} = \Delta T_{\text{gold}}$

(2) $\Delta T_{\text{plastic}} = 2 * \Delta T_{\text{gold}}$

(3) $\Delta T_{\text{plastic}} = \frac{1}{2} \Delta T_{\text{gold}}$

insufficient information

Figure 6 shows two photos of students discussing during the second round of a conceptTest. A variety of typical behaviors are visible. In the traditional theater (upper panel), students discuss with their immediate neighbors, but they also discuss with students in other rows. For example, a student toward the front right of the photo in the traditional lecture has leaned forward to discuss with students in the row in front of him. In the lecture theater with

the swivel seats (lower panel, Figure 6), some students choose a complete turn so they are face-to-face with another row, while others do a quarter turn so they face slightly inward toward the people in the same row (e.g., two of the groups in the back row of upper panel of Figure 6 have turned inward).

Results

Formative assessment

The software for the clicker system logs each student's answer, making it possible to track the improvement in cor-

rect answers between the two rounds of questions. Figure 7 shows a scatter plot of the percentage of correct answers during the second round as a function of the percentage correct during the first round. Eighty-three questions were included in this sample, which had data collected from all three lectures. The data from the traditional lecture theater (1:00 p.m.) are shown as blue filled squares, while those from the lecture theater with swivel seats are shown as red open circles for the 2:00 p.m. class, and red open triangles for the 3:00 p.m. class (also swivel seats).

The scores typically improve from the first to second round, though there are a few examples where this is not the case. It is also noted that once the first-round score reaches about 70% correct, then the second-round score often reaches close to 100%.

These results can be summarized by averaging the questions for the entire semester. Table 1 contains the averages and standard deviations for the percentage of correct answers for the two rounds. The second-round scores for the swivel lecture theater are approximately five percentage points higher than those

FIGURE 6

Student discussions during second round of questions. The top photo is the traditional lecture theater with one row per tier. On the bottom is the lecture theater with swivel seats and two rows per tier.



FIGURE 7

Scatter plot of percentage of students who got the correct answer during the second round versus the percentage of correct answers during the first round. The data from the traditional lecture theater (1:00 p.m.) are shown as blue squares, while those from the lecture theater with swivel seats are shown as red open circles for the 2:00 p.m. class, and red open triangles for the 3:00 p.m. class.

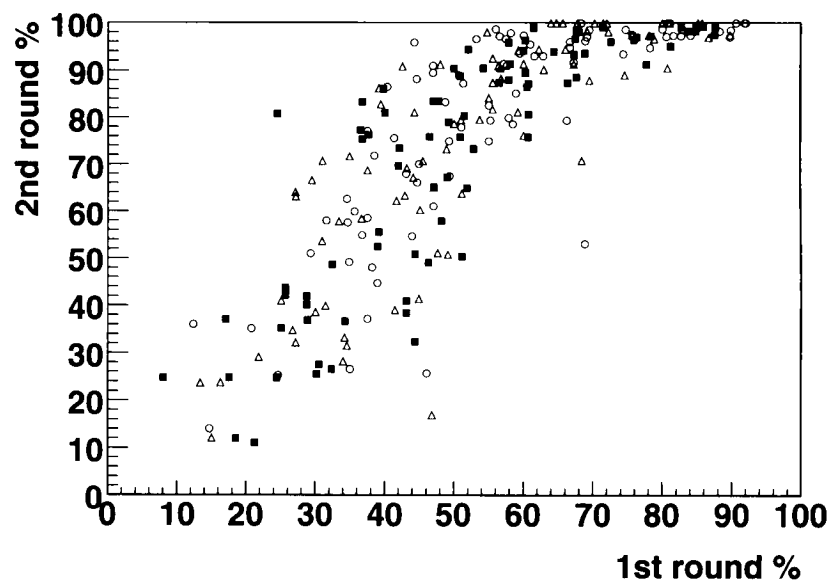


TABLE 1

Summary of conceptTest scores averaged over the 83 questions during the semester. One lecture was held in a traditional theater while the other two were held in a theater with swivel seats and two rows per tier. The results from the two swivel lectures are combined in the final column. The rows contain averages and standard deviations for the percentage of correct answers for the two rounds. The definitions of effect size and gain, g , are given in the text.

	Traditional (1:00 p.m.)	Swivel (2:00 p.m.)	Swivel (3:00 p.m.)	Swivel (combined)
<1 st round>	51.2 ±2.1	56.6 ±2.1	52.6 ±2.1	54.6 ±1.5
$\sigma_{1st\ round}$	19.1	18.8	18.9	19.0
<2 nd round>	71.9 ±2.8	79.7 ±2.4	74.2 ±2.7	76.9 ±1.8
$\sigma_{2nd\ round}$	25.7	22.0	24.6	23.5
Discussion effect size	1.08 ±0.18	1.22 ±0.17	1.14 ±0.18	1.18 ±0.12
< g >	0.524 ±0.038	0.626 ±0.037	0.552 ±0.038	0.589 ±0.027

from the traditional lecture theater, but it is noted that the first-round scores are also higher, a bias that will be explored later in this paper.

Table 1 also reports the pre- to postdiscussion effect size for each lecture group, where the discussion effect size is defined as

$$\text{discussion effect size} = \frac{\langle 2nd\ round \rangle - \langle 1st\ round \rangle}{\sigma_{1st\ round}}$$

The discussion effect sizes are larger for the swivel lecture theater, with the combined 2:00 p.m. and 3:00 p.m. classes larger than the traditional lecture theater effect size within one-sigma statistical significance.

Another measure that quantifies the improvement of scores during discussion is the relative gain (g) where for each question

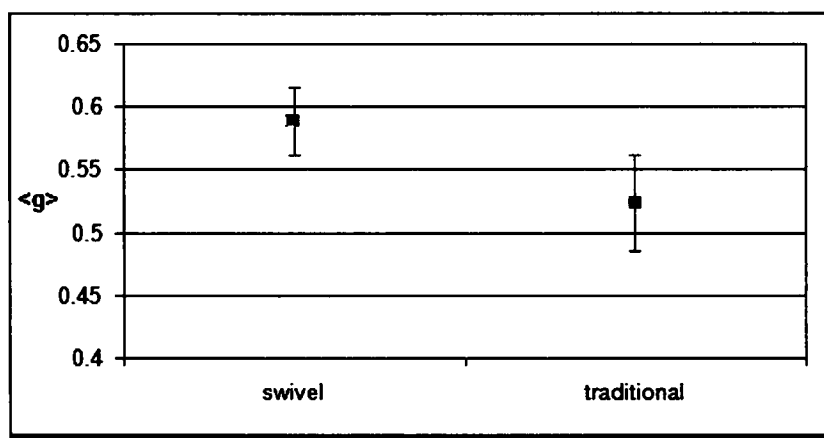
$$g = \frac{2nd\ round - 1st\ round}{100 - 1st\ round}$$

The gain is averaged over all 83 questions to produce $\langle g \rangle$ for each lecture group (Table 1; Figure 8). The average gains for the swivel lecture theater are higher than the traditional lecture, with the combined 2:00 p.m. and 3:00 p.m. classes larger than the traditional lecture theater at a statistical significance of $P = 0.002$ (two-tailed paired students' t -test. The pairing is made using the same questions in the two classes). However, this result has a potential bias because the swivel lecture groups had higher first-round scores.

The 83 questions can be binned into independent ranges of first-round scores, and the average second-round scores calculated for each interval. Five

FIGURE 8

Plot of $\langle g \rangle$ for traditional and swivel lecture theaters.



intervals were chosen ($0 < 1st\ round < 20\%$, $20\% < 1st\ round < 40\%$, $40\% < 1st\ round < 60\%$, $60\% < 1st\ round < 80\%$, and $80\% < 1st\ round < 100\%$). The average second-round scores for each interval are shown in Figure 9, where the x -coordinate of each data point is plotted at the average first-round value within that range.

In the second round, scores show very little difference between the two lecture theaters as a function of the first-round scores, suggesting that the difference in average gain observed

between the two lecture theaters is dominated by the slightly larger first-round scores in the swivel classes. There is a suggestion of a modest, one sigma, increase in second-round scores for the swivel lecture theater compared to the traditional theater when the first-round scores are close to 50%, i.e., when there is no clear-cut answer after the first round.

Summative assessment

While there is no immediate benefit to discussions in swivel seats compared

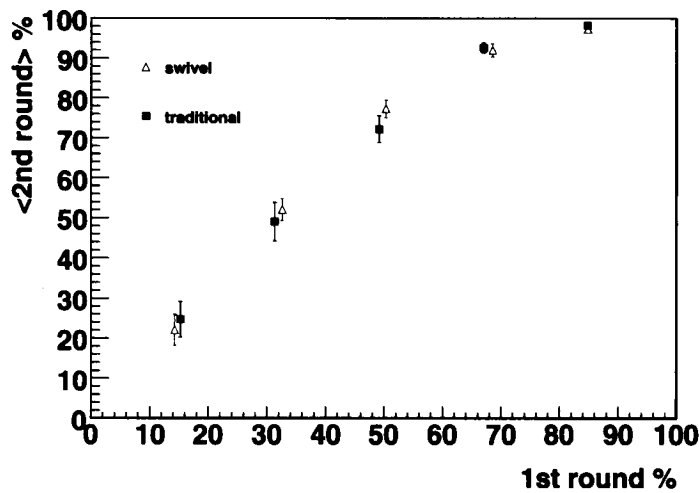
TABLE 2

List of average scores and variances for Test 1 (four weeks into semester) and for the final exam for students who attended at least half of the lectures in either the traditional or swivel-seat lecture theaters.

	Test 1 $\%$	Test 1 σ	Final $\%$	Final σ
Traditional	71.1 ±1.5	16.1	53.2 ±1.6	17.4
Swivel	73.5 ±1.3	14.2	58.9 ±1.6	17.0

FIGURE 9

Average second-round scores for different ranges of percentage correct answers during the first round. The data from the traditional lecture theater are shown as blue filled squares, while the data from the swivel lecture theater are shown as red open triangles.



in Table 3, where both the low- and high-performing groups of students who attended lectures in the swivel-seat theater scored higher on the final than their counterparts who attended lectures in the traditional theater.

These data can be characterized by calculating the swivel-seat effect size from the final scores,

$$\text{swivel - seat effect size} = \frac{\text{swivel} - \text{traditional}}{\sigma_{\text{traditional}}}$$

which is shown in Figure 10. Both high- and low-performing students appear to have benefited from the swivel-seat discussions by the end of the semester, with potentially a larger benefit for stronger students.

Conclusions

If an innovation in teaching is known to improve student learning, then to what extent do architecture and learning-space design enhance the effect of the technique? The example in this paper has been the use of cognitive conflict questions and peer discussions in large lectures, a technique that has been shown to improve student learning compared to students listening to a lecture for 50 minutes. Implementing swivel seats in a large lecture theater was an effort to enhance discussions among students and improve their understanding of the material.

The data were collected in the first semester after a major remodeling of two lecture theaters. One theater was a traditional design with

to traditional rows, it is worthwhile to explore whether there is any cumulative benefit over the whole semester. I do not have available a pretest for the course; instead, I use student scores from the first midterm exam, which was held four weeks into the semester, and compare those scores with student performance on the final exam. Table 2 lists the percentage correct scores for these two exams. Students were sorted into two groups by attending at least half of the lectures in either the traditional or swivel-seat lecture theaters. Both exams contained a mix of quantitative and conceptual questions: The final was clearly harder than the first test.

Students who attended lectures in the swivel-seat lecture theater scored almost six percentage points higher on the final exam than those in the traditional theater. These students also scored marginally higher at the first exam, so there is a potential bias of stronger students being in the swivel-lecture group. To reduce this bias, I further subdivided students into two groups: one group who scored less than 70% on the first test (70% is close to the average) and the other group who scored greater than 70% on the first exam. These subgroups of students

are well matched across the two lecture rooms. The averages for Test 1 for the low-performing group were 56.3% ±1.7% and 56.9% ±1.3% for the traditional and swivel-seat lecture theaters, respectively. For the high-performing group the corresponding averages for Test 1 were 81.9% ±1.7% and 81.8% ±1.3% for the traditional and swivel-seat lecture theaters respectively. Hence, this sub-grouping selects matched students across the two rooms.

The average scores for the final exam for these subgroups are listed

FIGURE 10

Swivel-seat effect size for the final scores between the two designs of lecture theaters. The data on the left are for all students, which are then split into two subgroups; those who were low performers on Test 1 and those who scored well on Test 1.

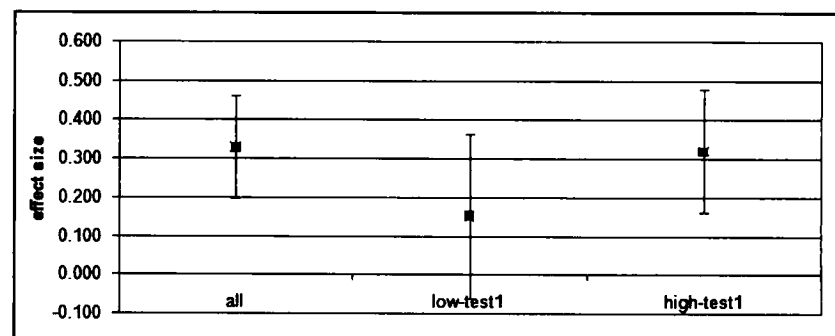


TABLE 3

List of scores for the final exam for students who attended at least half of the lectures in either the traditional or swivel-seat lecture theaters split into two subgroups: those that scored greater or lesser than 70% on Test 1.

	Final % traditional	Traditional variance	Final % swivel	Swivel variance
All	53.2 ± 1.6	17.4	58.9 ± 1.6	17.0
Test 1 <70%	39.6 ± 1.6	10.8	41.3 ± 1.7	10.6
Test 1 >70%	59.1 ± 2.3	18.3	64.9 ± 1.8	16.1

one row of seats per tier. The other theater had two rows of seats per tier, and the front row could swivel so that students could face each other during discussions.

Students in the swivel lecture theater achieved higher scores after discussion than students in the traditional lecture theater. However, most, if not all, of this difference could be attributable to students in the swivel classes achieving a slightly larger score on the first round of questions. One avenue of future work would be to stop displaying the frequency of the first-round answers to students before the second-round discussion. This would remove some of the pegging effect observed when one answer dominates the first round, but it could also reduce the level of engagement that occurs when the histogram reveals that there is no clear-cut preferred answer during the first round.

I also observed that students who attended lectures in the swivel-seat theater scored higher on the final exam. This effect persisted when I divided both lecture groups into high- and low-performing groups based on their test scores early in the semester. Both high- and low-performing students appear to have benefited from the swivel-seat discussions by the end of the semester, with potentially a larger benefit for stronger students. ■

References

Bartlett, A.A., and M.A. Thomason. 1983. Legibility in the lecture hall. *Physics Teacher* 21 (8): 531.
Bartlett, T. 2003. Take my chair (please). *Chronicle of Higher Edu-*

cation 49 (26): 36–38.

Bransford, J.D., A.L. Brown, and R.C. Cocking, eds. 2000. *How people learn*. Washington, DC: National Academy Press.

Dane, J. 2004. Designing environments that stimulate student-centered learning. www.tefma.com/infoservices/papers/2004_Hobart_Annual_Conference/Dane.pdf.

Douglas, D., and R. Gifford. 2001. Evaluation of the physical classroom by students and professors: A lens model approach. *Educational Research* 43 (3): 295–309.

Fleming, D., and J. Starr. 1999. The impact of lecture theatre design on learning experience. *Facilities* 17 (7/8): 231–36.

Halloun, I., and D. Hestenes. 1985. Common sense concepts about motion. *American Journal of Physics* 53 (11): 1056–65.

Heller, P., and M. Hollabaugh. 1992. Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics* 60 (7): 637–644.

Heller, P., R. Keith, and S. Anderson. 1992. Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics* 60 (7): 627–36.

Jamieson, P., J. Dane, and P.C. Lippman. 2005. Moving beyond the classroom: Accommodating the changing pedagogy of higher education. Proceedings of the 2005 Forum of the Australasian Association for Institutional Research, 17.

Johnson, D.W., R.T. Johnson, and K.A. Smith. 1991. *Cooperative learning:*

Increasing college faculty instructional productivity. Washington, DC: George Washington University, School of Education and Human Development.

Lang, D. 1996. Essential criteria for an ideal learning environment. *New Horizons for Learning: Teaching and Learning Strategies* 2 (1).

Mazur, E. 1997. *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.

McDermott, L.C., and P.S. Shaffer. 1992. Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics* 60 (11): 994.

Oblinger, D. 2005. Leading the transition from classrooms to learning spaces. *EDUCAUSE Quarterly* 28: 1.

Piaget, J. 1985. *The equilibration of cognitive structure: The central problem of intellectual development*. Chicago: The University of Chicago Press.

Project Kaleidscope. 1998. What difference do improved facilities make. www.pkal.org/keywords/Facilities.cfm.

Redish, E., J. Saul, and R. Steinberg. 1998. Student expectations in introductory physics. *American Journal of Physics* 66 (3): 212–24.

Shaffer, P.S., and L.C. McDermott. 1992. Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of instructional strategies. *American Journal of Physics* 60 (11): 1003.

Strange, C.C., and J.H. Banning. 2001. *Education by design*. San Francisco: Jossey-Bass.

University of Maryland, Baltimore County (UMBC). 2000. Lecture hall design standards. www.umbc.edu/oit/classroomtechnology/av/avhome/LH_Standard.pdf.

Craig A. Ogilvie (cogilvie@iastate.edu) is a professor of physics at Iowa State University in Ames, Iowa.