Chapter 9

Experimental and Quasi-Experimental Research

9.1 INTRODUCTION

Research on the performance of various building components constitutes a significant and longstanding domain within architectural research as a whole. Although much of this research has focused on improving building technologies in the industrialized world, a study by Givoni, Gulich, and Gomez focuses instead on radiant cooling by metal roofs, a significant issue for housing in developing countries.[1] Givoni et al. noted that although corrugated metal roofs are effective for cooling in the evening, they tend to overheat houses in the daylight hours. The researchers hypothesized that the installation of operable hinged interior insulating plates under the roof would reduce daytime heating without interfering with the nighttime cooling function of the metal roofs.

To test this hypothesis the researchers built a small-scale mock-up of the typical house (termed a *test cell*) whereby the heating/cooling effect of various test conditions could be measured. (See Figure 9.1.) Givoni et al. tested three distinct conditions of insulation operation: 1) with the insulation panels closed both day and night; 2) with the insulation panels open at night and closed during the day; and 3) with the insulation positioned as in condition 2, but with the addition of a small ventilating fan from midnight to 5 am. In addition, two levels of thermal mass (as represented by water-filled bottles) were also tested.

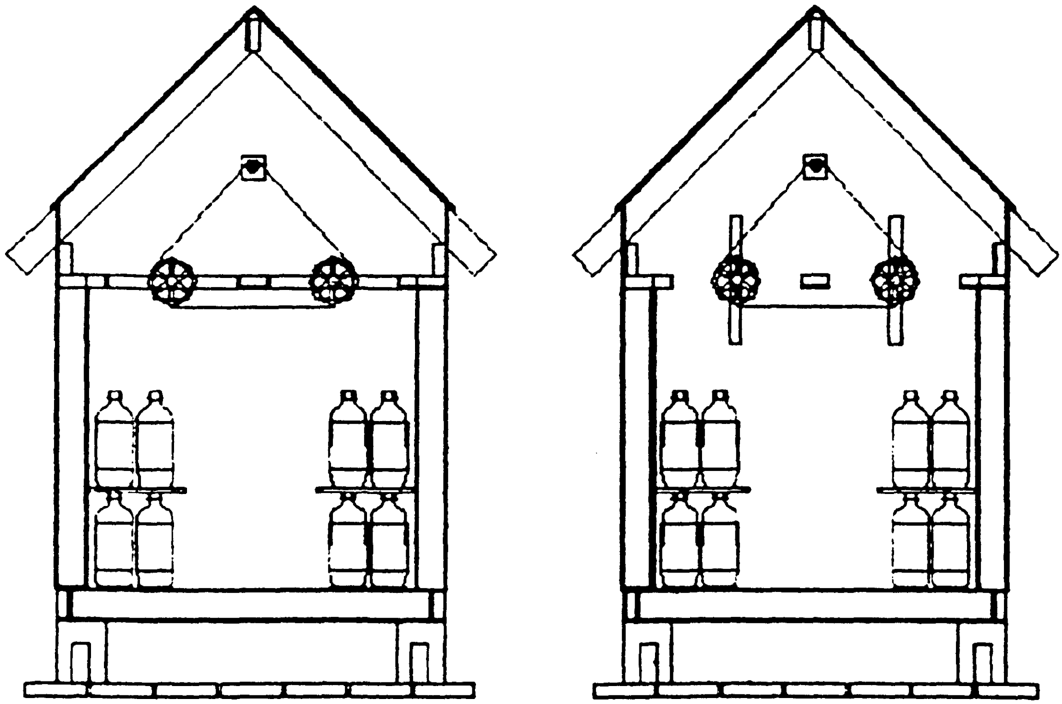

Figure 9.1 Test cell unit, showing panels closed (left) and open (right), Baruch Givoni et al., 1996. Courtesy of American Solar Energy Society, Inc.

Based on their tests of these conditions, the authors concluded that the combination of insulating panels and fan venting (condition 3) provides better daytime cooling than the insulation without the fan ventilation. On the other hand, no appreciable difference in cooling was noted as a consequence of the thermal mass condition. Finally, based on these data, the authors were able to develop predictive formulae for calculating the indoor maximum temperature as a function of the swing of the outdoor temperature.

Taking on a very different research topic, Ann Sloan Devlin sought to discover the extent to which gender might have an effect on how job applicants are evaluated in architectural practice.[2] She hypothesized that "women architects would be less favorably rated than male architects," especially at the more senior level.[3]

To test this hypothesis, Devlin created both a junior-level and senior-level resume, the junior level with four years of architectural experience and the senior level with 13 years of experience. Copies of each resume type (junior and senior) were created using a fictitious female name; an equal number of copies carried a fictitious male name. Each resume included a career objective, professional experience, affiliation, registration, education, skills, honors and awards.

Respondents in the study were more than 200 architects (156 men and 48 women) licensed in the state of Connecticut, but representing all regions of the country. Respondents were told that the study was about "the perception architects have of the characteristics possessed by those practicing architecture." These respondents then received one of the four fictitious resumes and were asked to evaluate the candidates on a seven-point scale in the following areas: technical aspects of the job, administrative aspects, interpersonal aspects, contribution to growth of firm's client base, creative contribution, advancement, and overall ability. Respondents were also asked whether they would accept or reject the candidate for hire.

The most salient result of Devlin's study was that the "male architect respondents were more likely to hire male applicants than female applicants as senior architects."[4] Devlin reached this conclusion by comparing the hiring decisions of the respondents in relation to the four resume conditions (male or female; intern or senior), using inferential statistical measures (see Chapter 8, section 8.3.1). She concludes that women in architecture may indeed "experience discrimination as they advance through the ranks."[5]

9.2 STRATEGY: GENERAL CHARACTERISTICS OF EXPERIMENTAL RESEARCH

These two studies may seem to be worlds apart. On a thematic level, the Givoni et al. study tackles an aspect of environmental technology, while the Devlin study seeks to clarify the dynamics of gender discrimination in architectural practice. The research contexts are also very different. The former is conducted in a laboratory setting, while the latter makes use of a real-life or "field" setting. The variables being investigated are quite different as well. The Givoni et al. study considers only physical variables; whereas the Devlin study focuses on behavioral and social conditions.

Despite these notable differences, the Givoni et al. and Devlin studies are alike in that they are both examples of the experimental research design. Many readers are likely to read into that factual statement either a commendation of high praise or an invitation to criticism. This is because experimental research is so frequently portrayed as the standard against which all other research strategies should be judged. In general, readers who adhere to the postpositivist system of inquiry are likely to see the experimental strategy as the essence of "scientific" research. On the other hand, many researchers who adhere to a naturalistic or emancipatory paradigm have argued persuasively that the experimental design is either inappropriate or insufficient for research about certain social and cultural phenomena. We will address some of these concerns later in this chapter. (See section 9.6.) We would argue that, as with the other research strategies, experimental research can yield either outstanding or flawed research, depending on how appropriately it is applied to a particular research question.

What then are the underlying commonalities that define the Givoni et al. and Devlin studies as experimental research? Briefly, the defining characteristics of an experimental research design include the following: the use of a treatment, or independent variable; the measurement of outcome, or dependent, variables; a clear unit of assignment (to the treatment); the use of a comparison (or control) group; and a focus on causality.[6] These five characteristics will be discussed in some detail in the following chapter segments.

9.2.1 *The Use of a Treatment, or Independent Variable*

In each of the two studies described above, the researchers are seeking to study the impact of one or more specific, identifiable variables on the phenomenon under study. In the case of the metal roof study, the researchers are seeking to test the thermal impact of several conditions, both in isolation and in combination, including: *insulation*, *venting fan*, and *thermal mass*. Similarly, in her research on gender issues in professional practice, Ann Sloan Devlin is seeking to clarify the impact of *gender designations* on how architects evaluate job applicants. Although quite different in nature, these variables are in each case *manipulated* or *controlled* by the researchers in some specified way, and so they are considered to be *treatments* in the experimental strategy.

9.2.2 *The Measurement of One or More Outcome Variables*

In each of these studies, the researchers were able to specify the impact of the experimental treatment by carefully measuring certain outcome measures, or dependent variables. For Givoni et al.'s study of metal roofs, the dependent variables were the *temperature readings for indoor areas* of the test cell environments including both the attic and the indoor living environment. The researchers were able to ascertain how much the indoor temperatures were cooled by the several experimental conditions (see Figures 9.2 and 9.3). In a similar way, Devlin was able to assess the impact of gender designations through two measures: a questionnaire instrument whereby prospective employers could register their *evaluation* on a 1-to-7 rating scale, and a *hiring decision* to accept or reject. Again, although quite different in nature, both the temperature and evaluation measures are the outcome measures (or dependent variables) of these experiments.

9.2.3 *The Designation of a Unit of Assignment*

In each of these studies, the researchers have applied the experimental treatment to a specified *unit of assignment*. In the case of Givoni et al.'s research, the treatment conditions (various combinations of insulation, venting fans, and mass) are all applied to a *test cell*. This test cell was a small-scale mock-up of a metal-roofed residential unit in a hot climate, a 1-meter cube with metal gabled roof. (See Figure 9.1.) On the other hand, in Devlin's study the unit of assignment was not an inanimate object, but rather the *individual architects* who were asked to evaluate the fictitious job applicants. Each of these "units"—whether test cells or individual architects—received a treatment manipulated by the researcher(s).

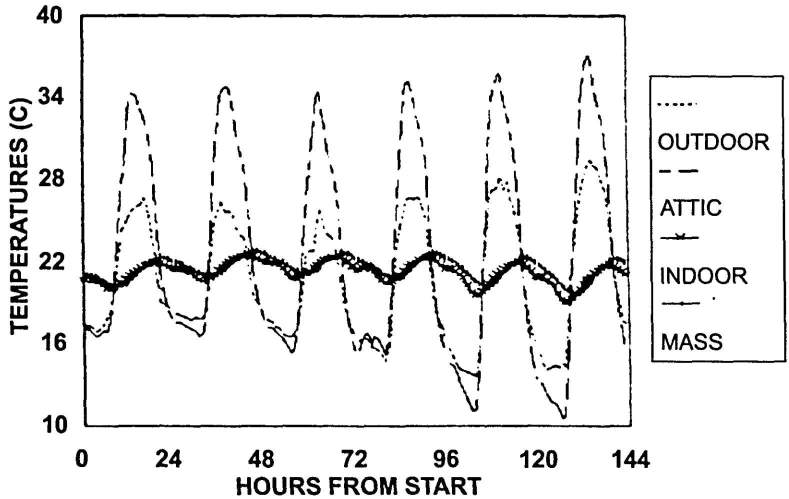

Figure 9.2 Control conditions: insulating panels closed day and night. Courtesy of American Solar Energy Society, Inc.

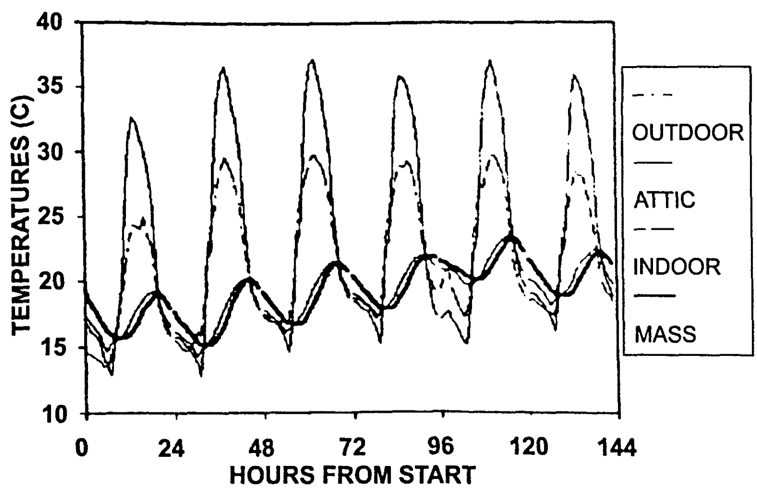

Figure 9.3 Insulating panels closed during the day and open at night. Courtesy of American Solar Energy Society, Inc.

9.2.4 *The Use of a Comparison or Control Group*

Most experimental studies measure the impact of treatments against a comparison or control group. The control condition in Givoni et al.'s study is achieved with the insulation panels closed both day and night so that no heating or cooling occurs. In all other conditions, i.e. the treatment conditions, the insulation panels are closed during the day and opened at night to allow for cooling. In other words, the control condition is defined as one to which the treatment is *not* applied. In Devlin's study, all architect respondents received some treatment, one of four combinations of male or female applicant, at a junior or senior level. In this case, the different treatments are compared against each other.

9.2.5 *A Focus on Causality*

The combined purpose of the defining features of the experimental research design (i.e. treatment, outcome measures, unit of assignment, and control or comparison groups) is to enable the researcher to credibly establish a cause-effect relationship. In general, the experimental researcher is seeking to ascertain and measure the extent to which a treatment causes a clearly measured outcome within a specified research setting, whether in a laboratory or in the field.

Although the underlying structure of the experimental research design is essentially consistent across diverse topic areas, researchers vary in the extent to which they take "causality" for granted.[7] Experimental research in environmental technology (such as the metal roof study) is more likely to take causality for granted than research in socio-cultural aspects of architecture (such as the gender designation study). This is because environmental technology, like much research in natural science, tends to incorporate the following characteristics: 1) the use of laboratory settings where relevant variables can be easily controlled; 2) dependent variables that are in many instances inert, and therefore not likely to change except as a consequence of the treatment; 3) explicit theories that enable researchers to specify the expected effects of a particular treatment; and 4) instruments that are calibrated to measure such effects. Given these conditions, causality can often be assumed without much discussion or argument.

In research that involves people's reactions or behaviors, or more complex arrays of variables in field settings (as in Devlin's research), researchers tend to be more explicit about how they have met the basic requirements of experimental design. Likewise, in drawing their conclusions, researchers who explore socio-physical dynamics in architecture tend to emphasize the conditions and limitations of a causal interpretation. Devlin, for example, qualifies her conclusion that male respondents tend to rate senior female applicants less positively than senior male applicants. Devlin mentions two limitations to a causal interpretation: 1) many respondents explained that they found it hard to rate the applicants because the resume information was so limited; and 2) the response rate was only 30% and therefore the extent of generalizability to the larger population of architect employers is limited. Such problems and limitations in experimental research will be discussed in greater detail in segment 9.5 of this chapter.

9.3 STRATEGY: DISTINGUISHING BETWEEN EXPERIMENTAL AND QUASI-EXPERIMENTAL RESEARCH

So far in our discussion, we have discussed only the general requirements of experimental research, without recognizing the very important distinction between experimental and quasi-experimental designs. This distinction rests on the manner in which the units of assignment (whether test cells, people, etc.) are selected. Although the goal for both experimental and quasi-experimental research is to achieve comparability among the units in each treatment group, such comparability is more precisely established in experimental research through random assignment. In contrast, the quasi-experimental research design is often employed in field settings where people or groups cannot be randomly assigned for either ethical or practical reasons. In such cases, the researcher seeks to ascertain or establish effective comparability across as many variables as possible. These considerations are discussed in greater detail below.

9.3.1 Random Assignment in Experimental Research

Random assignment plays an important role in experimental research when there is reason to believe that the units of assignment are not completely equivalent. It is considered the most effective way to ensure the essential comparability of treatment groups so that the observed differences in outcome measurements can be credibly attributed to the treatment.

In the gender discrimination study, Devlin was able to employ random assignment, even though the respondents were not conducting their evaluations in a laboratory setting. By choosing to manipulate the resume conditions rather than depend on the real-life applicant resumes received by these architects, Devlin could randomly assign the architects (registered in Connecticut, Devlin's home state) to the various resume conditions. This provides a greater level of assurance that the gender of the applicant actually had an effect on the male architects' evaluations.

On the other hand, experimental research based on inert materials (such as the Givoni et al. study) does not necessarily require such randomization measures. In most circumstances, the essential comparability of test cells or mock-ups can be assumed either because: 1) materials of the same physical specifications are used; or 2) the same physical unit can be reused in a different treatment condition. As a consequence, the authors of the metal roof study can claim that, given certain specified climatic conditions, the different measured cooling outcomes can be attributed to the differences in treatment conditions.

9.3.2 Nonrandom Assignment in Quasi-Experimental Research

As mentioned earlier, research studies conducted in the field frequently entail situations in which random assignments cannot be achieved for either ethical or practical reasons. For example, if a researcher wanted to test the effect of four lighting systems on employee productivity in four separate office areas, it is unlikely that managers would agree to assign their employees randomly to the four office areas in a way that would disrupt important work group functions.

In this situation, researchers would likely adopt a quasi-experimental design in which they would identify four *existing* work groups, each of which would receive a different lighting treatment. In doing so, the researchers would attempt to find work groups comparable in as many respects as possible, including their task or work objectives, mix of job types, gender mix, age range, level of education, etc. If, for instance, the work groups' tasks were quite dissimilar, it would then be more difficult to attribute measured differences in productivity to the lighting treatment rather than differences in the tasks.

Another example of quasi-experimental design is a small research project conceived and conducted by students in one of Groat's research methods classes.[8] The students had raised the issue of a small gallery area near the school offices that had been created to function as both an exhibit space and a lounge area for faculty and students. The students observed that the space was seldom used as a lounge. Discussion soon revolved around what sort of changes would have to made for the area to function more as a lounge and social space. The students hypothesized that the gallery would be used more if the arrangement of furniture were less formal and if small screening elements were used to block the view through the glass wall along the doorway side of the space.

The students' research design involved two sets of observations of the space. The first observations recorded people's use of the space in its existing condition; and the second recorded its use under the experimental treatment. The observations were made on the Monday (studio day) and Tuesday (nonstudio day) of two successive weeks, starting at 8:30 in the morning and continuing to 7:30 at night. Each observation period was of 15 minutes duration, starting on the half-hour and ending at 45 minutes after each hour.

The experimental treatment condition, used in the second two-day observation period, was designed to create a more "inviting" ambience; it entailed alteration of the furniture arrangement, lighting levels, and ambient sound. (See Figures 9.4 and 9.5.) More specifically, the following alterations were made: addition of screening elements to create more visual privacy from the hallway windows; relocation of some furniture elements for more privacy and to create groupings; lowering of fluorescent lighting

Figure 9.4 Existing condition of the gallery space, from Janice Barnes et al. Photo courtesy of Barnes et al.

Figure 9.5 Modified condition of the gallery place. Photo courtesy of Barnes et al.

levels; addition of incandescent table lamps; introduction of reading materials on the tables; use of soft background music; and introduction of plants.

The students also developed a one-page observation sheet that included the following information: a count of the number of people using the space during that observation period; a plan of the gallery including the furniture arrangement in which the people's movement and activities were mapped; and a coding system by which people's specific activities could be described (i.e. speaking, writing/reading, sleeping).

The general conclusion that the students were able to draw was that although the numbers of people using the space did not change substantially, the average amount of time each person spent in the gallery increased, and the nature of their activities changed as well. (See Figures 9.6–9.11.) Indeed, by the second day of the treatment condition, the proportion of staying activities was more than double that of the previous Tuesday in the control condition.

How much of this change can be attributed to the treatment effect? The circumstances of the field setting did not allow the students to assign gallery users randomly

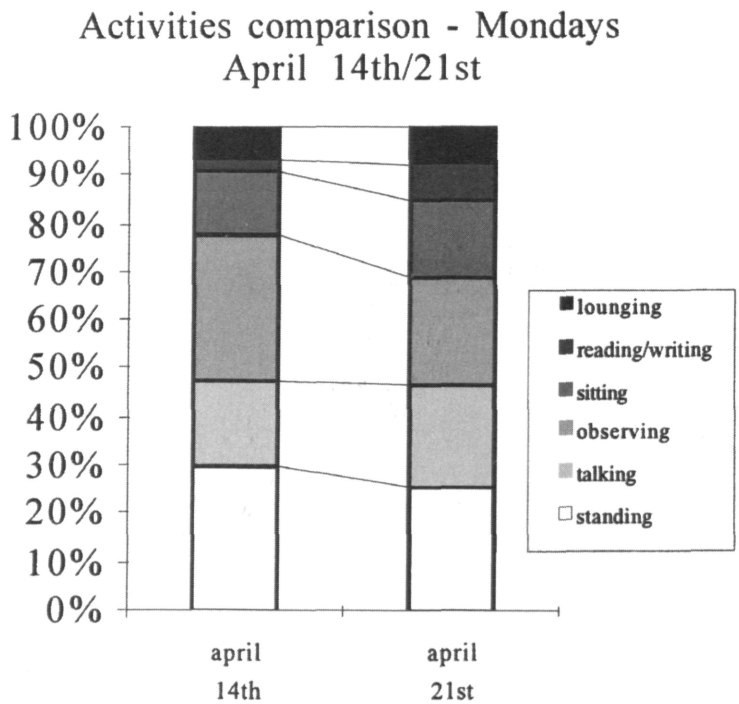

Figure 9.6 Comparison of observed activities for the existing and modified conditions, Mondays. Courtesy of Barnes et al.

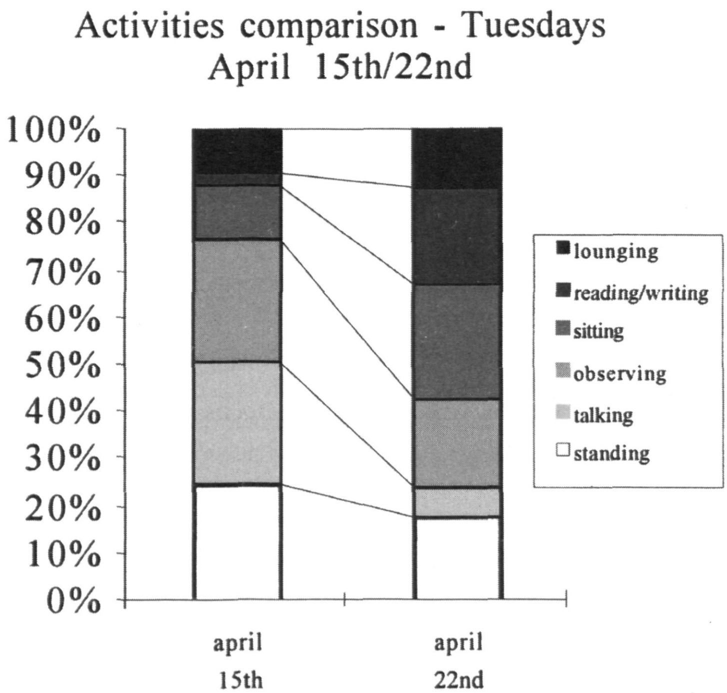

Figure 9.7 Comparison of observed activities for the existing and modified conditions, Tuesday. Courtesy of Barnes et al.

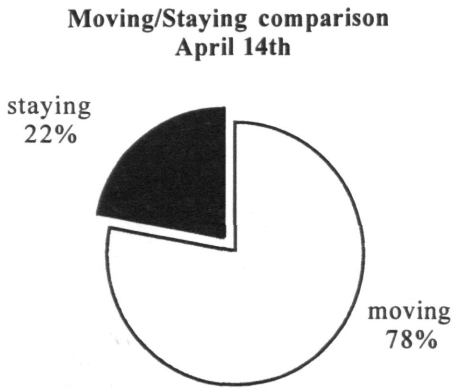

Figure 9.8 Moving/staying activities for the existing condition, Monday. Courtesy of Barnes et al.

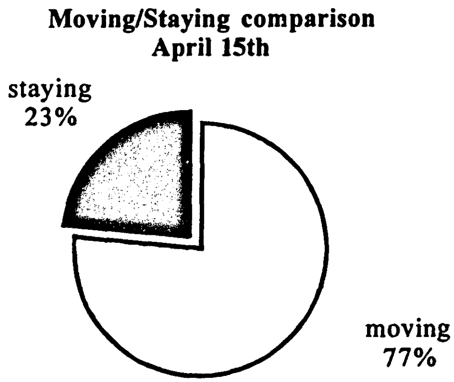

Figure 9.9 Moving/staying activities for the existing condition, Tuesday. Courtesy of Barnes et al.

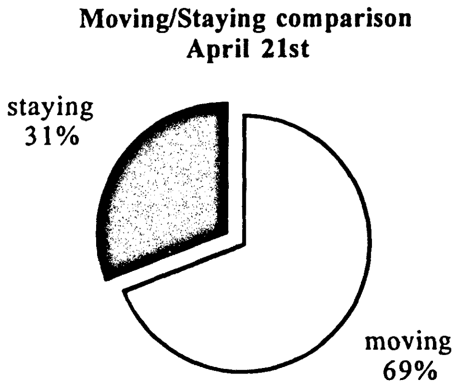

Figure 9.10 Moving/staying activities for the modified condition, Monday. Courtesy of Barnes et al.

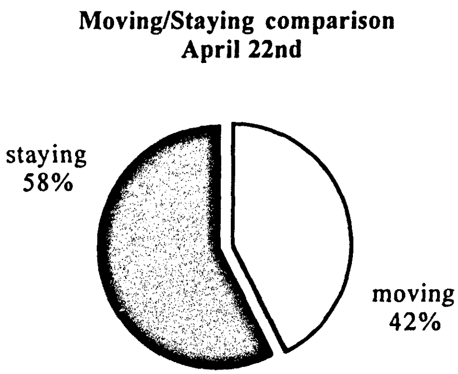

Figure 9.11 Moving/staying activities for the modified condition, Tuesday. Courtesy of Barnes et al.

to the two conditions, and so they adopted a quasi-experimental design. Since no specific measures of the gallery users were taken, it is not possible to gauge precisely how the users of the control condition compared with those in the treatment condition. Still, there were no obvious indicators that the groups were substantially nonequivalent. It is therefore likely, but not certain, that the "informal, inviting" condition did encourage a change in the use patterns of the gallery space.[9]

9.4 DIAGRAMMING EXPERIMENTAL RESEARCH DESIGNS

From the experience of the architectural design process, we know that it is often helpful, sometimes even essential, to diagram the singular qualities of a design concept or parti. In a similar vein, experimental researchers have devised a way of diagramming the particular details of experimental research designs, using the following coding system:

{R = Random assignment}

{X = Experimental treatment}

{O = Observation of dependent variables (e.g., pretest or posttest)}

Although there are a great many standard experimental research designs that use an established nomenclature,[10] we will limit our discussion to the three exemplar studies that have been discussed thus far in the chapter.

The Givoni et al. study of radiant cooling is represented below. Each row represents, from left to right, the sequence entailed in each treatment condition.

	O	{Observation only, with no prior treatment}
X1	O	{Treatment 1, and subsequent observation}
X2	O	{Treatment 2, and subsequent observation}
X3	O	{Treatment 3, and subsequent observation}

This notation system conveys the following essential points about the design of this study: 1) there is no explicit attention paid to random assignment, since all the relevant procedures deal with standardized inert materials; 2) there are three different treatment conditions in addition to the control condition; and 3) only posttest (i.e. no pretest) observations are made.

Devlin's study of gender issues in architectural practice presents a slightly different research design in the following respects: 1) random assignment is an explicit and important consideration for establishing comparability across treatment groups; and 2) there is no explicit control condition. However, as in Givoni et al.'s study, no pretest observations are made. Thus the notation system for this study can be represented this way:

R X1 O	{Random assignment, followed by treatment 1, observation}
R X2 O	{Random assignment, followed by treatment 2, observation}
R X3 O	{Random assignment, followed by treatment 3, observation}
R X4 O	{Random assignment, followed by treatment 4, observation}

The student study of behavioral patterns in a gallery space presents a slightly more ambiguous research design. This is because the researchers did not clarify the extent to which the people who experienced the original gallery arrangement were the same people who experienced the modified arrangement. (This could have been achieved by asking users if they had come into the gallery anytime during the previous Monday or Tuesday.) If the gallery users had been substantially the same group, then the notation of the research design would be as follows:

O O X O O	{Two observations, treatment, followed by two observations}

This design is known as a single-group interrupted time-series design. Two pretest observations were made, after which the treatment (physical modification) was applied, followed by two posttest observations.

On the other hand, if the two sets of users were substantially or completely different, then it would be more accurate to diagram the research design in the following way:

O O	{No treatment, two observations only}
X O O	{Treatment, followed by two observations}

This second diagram presumes that the group that experienced the original gallery arrangement constitutes the control group, whereas the group that experienced the new arrangement was the experimental treatment group. Both control and treatment groups were observed twice, the treatment group only as a posttest.

Readers who choose to make use of experimental research procedures are advised to consult some of the books referenced at the end of this chapter for further examples of experimental designs. These diagrammatic notations can be exceedingly useful to the researcher for clarifying the precise nature and assumptions of the selected experimental design.

9.5 TACTICS: THE SETTINGS, TREATMENTS, AND MEASURES FOR EXPERIMENTAL RESEARCH

Experimental research can involve a wide variety of tactics. The experimental setting can range from a highly controlled laboratory to less well-controlled field site. Similarly, the treatment conditions can range from highly calibrated physical manipulations to categorical, nonphysical conditions, such as the gender designations in Devlin's study. And finally, measurement of the outcome variables can range from the precise calibration of a physical change (such as that of air temperature in the Givoni et al. study) to the more descriptive index of a behavioral response (such as in Devlin's study).

In the sections that follow, the broad range and combinations of tactics available to experimental and quasi-experimental research will be discussed in the context of several specific research studies.

9.5.1 Tactics Used in the Example Studies

Before considering additional examples of experimental research, we would like to characterize more explicitly the tactics used in the studies cited above. For instance, Givoni et al.'s study of radiant cooling employs the sort of tactics typically associated with experimental research in environmental technology. The construction and treatment of the test cells was carefully monitored within a university lab setting. The physical treatment conditions of the test cells could be precisely specified and controlled by the experimenters; and likewise the outcome measures of air temperature could be exactly measured by laboratory instruments. (See Figure 9.12 for a complete summary of the tactics used in the experimental studies cited earlier in this chapter. See Figure 9.16 for a summary of tactics used in studies cited in the remainder of the chapter.)

Study	Setting	Treatment	Outcome Measures
1. Radiant cooling (Givoni et al.)	Lab	Environmental modifications Modification insulation venting mass	Instrumented measures air temperature
2. Gender issues (Devlin)	Field	Resumes gender and seniority	Attitudinal response applicant evaluation hiring decision
3. Gallery behavior (Barnes et al.)	Field	Environmental modifications furniture lighting ambient sound	Behavioral change staying/moving screens

Figure 9.12 Summary of tactics in cited studies.

The Devlin study represents a set of experimental procedures starkly different from the Givoni et al. study. Indeed, one could argue that the combination of the setting, treatments, and measures in Devlin's study places it at the opposite end of the spectrum. First, the research setting is not only a field setting, but one that is dispersed across the country, in the many offices where the architects received the resume. Second, although the treatment conditions were conveyed physically in print through gendered names and stated levels of employment experience, the physical and interactive reality of a real-life applicant was absent. And finally, the outcome measures of evaluation and employment decision were rendered through scores on a questionnaire. In all of these ways, the focus of the study was on the social-cultural implications of non-physical treatment conditions, measured through attitudinal responses.

The student study of the architecture gallery, though quasi-experimental in design, represents an intermediate range of tactics. Although the study employs a field setting rather than a lab, this setting is relatively small and easily manipulated by the researchers. The treatment conditions are all physically based (e.g., arrangement of furniture, the type of lighting); they can be clearly specified and measured. And although the outcome is behavioral and requires some interpretation, the standards for counting people and classifying behavior are easily established.

9.5.2 Occupant Comfort from Air Movement: Using a Lab Setting, Physical Treatments, Instrumentation, and Subjective Measures

Although much environmental technology research relies on combining lab settings with exclusively instrumented measures of physical outcome variables, many other variations of lab setting research are possible. One example is a study by Edward Arens et al. concerning the use of personally controlled air fans to achieve cooling and perceived comfort.[11] The goal of this study was to evaluate the effectiveness of using fans, instead of compressor-based air conditioning, to achieve cooling comfort. The study was conducted in an environmental chamber (i.e. lab setting) where individual subjects could be exposed to a controlled range of warm temperatures. (See Figure 9-13.) The environmental chamber was designed to "appear as a realistic residential or office space."[12]

The 119 subjects (57 female, 62 male) were divided into two comparison groups. One group was asked to control the fan settings "in a naturally fluctuating outdoor mode"; and the second group used the fan's constant mode, "in which the inherent turbulence of the airstream was at higher frequencies than in the fluctuating mode."[13] During both experimental protocols, the subjects' time in the experimental chamber included two distinct activity segments generating two distinct metabolic rates: one that included both sitting and step-climbing (1.2 met), and another that was entirely sedentary (1.0 met). Throughout all sessions, the subjects experienced a range of temperatures from 25 to 30 degrees C. Thus the treatments represented a combination of both lab-based controls and behavioral regimens.

Figure 9.13 Environmental chamber. Photo by Marc Fountain, courtesy of Prof. Edward Arens.

The outcome measures included both instrumentation and subjective ratings of comfort. The former involved recording the subject's choice of fan speed; the latter used a seven-point scale from cold to hot indicating how the subject experienced the temperature of the environment. More than 80% of the subjects at the 1.2 met condition were able to maintain comfort up to 29 degrees C. As a result, the researchers were able to conclude that within certain temperature zones, the use of personal air fans can serve as an effective alternative to mechanical air-conditioning.

9.5.3 Experimental Validation of Simulation: Using a Lab Setting, Physical Treatments, and Instrumented Measures

Environmental technology research sometimes employs a combination of experimental and simulation strategies. Medved and Novak's study of double-pane windows, described briefly in Chapter 2, is an example of this combined strategy.[14] The researchers' objective was to evaluate the effectiveness of a double-pane window design with a screen and a siphon forming a semi-open cavity. (See Figure 2.1.) They used a mathematical and numerical modeling (simulation) strategy, which they then validated through experimentation.

BOX 9.1

Experiment: Energy Conservation in Housing

Malcolm Bell and Robert Lowe sought to test the impact of various energy-saving techniques in housing administered by the Housing Authority of York, United Kingdom.* (See Figure 9.14.)

In this field-setting experiment, the authors measured the impact of energy-saving improvements in modernized housing against a "control group of dwellings in the same modernization scheme but with no additional energy efficiency works."** The 21 houses in the experimental group were modernized with a combination of clearly specified physical treatments: insulation, draft-proofing of doors and windows, central heating with gas-condensing boiler, and a gas fire as a secondary heat source. The 11 houses in the control group, with no additional energy efficiency works, were well matched with the experimental houses in terms of their initial energy consumption. As a consequence, any consistent differences in energy consumption could be attributed to the experimental treatment.

Figure 9.14 Typical house type in Malcolm Bell and Robert Lowe's energy-efficient modernization study. Reprinted from *Energy and Buildings* 32 (2000), with permission from Elsevier Science.

Monitoring measures included internal temperatures and gross energy consumption for the entire period, both of which were based on instrumentation. Although the difference of 5536 kwh between the experimental and control groups is statistically significant at the .03 level, the measured savings are about half of what was predicted by energy modeling. Further investigation, including interviews with residents, indicated that some residents used the secondary heat source, the gas fire, so often that the energy efficiency of the gas boiler was compromised. The monitoring of energy-efficient modifications in this real-world housing setting thus provided important insights about the limits of conservation hardware when not accompanied by changes in human behavior.

---

*Malcolm Bell and Robert Lowe, "Energy Efficient Modernization of Housing: A UK Case Study," *Energy and Buildings* 32 (2000): 267–280.

**Ibid., 272.

The experimental design, in this case, involved the use of a laboratory setting, physical construction of the treatment condition, and instrumented measurement of the outcome variables. The lab setting was a temperature-controlled room into which a large "hot box" was placed; the hot box was a cube, on one side of which the double-pane window design was installed. In the first series of experimental treatments, the temperature inside the hot box was higher than in the controlled chamber; whereas in the second series of experiments, the air inside the hot box was cooled. Outcome measurements included the following: thermacouples measured air and surface temperatures; bulb thermometers measured radiation temperature; and flux sensors measured heat fluxes. (See Figure 9.15.)

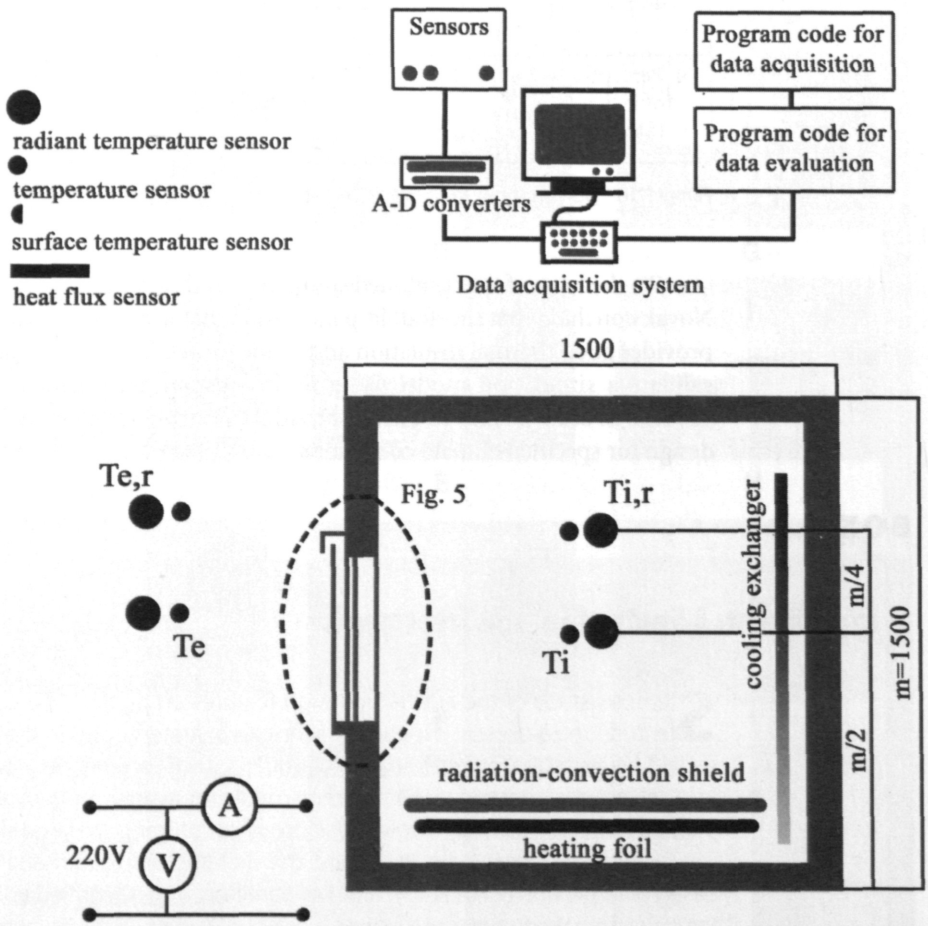

Figure 9.15 Experimental device by Saso Medved and Peter Novak. Reprinted from *Energy and Buildings* 28 (1998), with permission from Elsevier Science.

1. Personally controlled air fans (Arens et al.)	Lab	Physical treatments temperature activity level fan type	Instrumented measures and behavioral response fan speed choice perceived comfort
2. Energy use in housing (Bell and Lowe)	Field	Environmental modifications gas boiler insulation draft proofing secondary heat	Instrumented measures internal temperature gross energy consumption
3. Window pane design (Medved and Novak)	Lab	Environmental change hot box temperature radiation temp heat fluxes	Instrumented measures temperature measures
4. Perceptions of facades (Stamps)	Lab	Treatment of facade features	Perception of architectural mass

Figure 9.16 Summary of tactics in cited studies.

On the basis of the combined simulation and experimental strategy, Medved and Novak conclude that the double-pane design that includes a cavity with a "y" siphon provides good thermal insulation and shade protection. The researchers were able to validate a simulation model using limited experimentation; and that simulation model was used to achieve a much broader performance analysis of the windowpane design for specified climate conditions.

BOX 9.2

Experiment: A Study of Facade Treatments

Stamps's study of the effects of design features on people's perceptions of architectural mass uses an experimental research strategy, and in that regard it is unusual.* Many, probably most, studies of nonarchitects' or users' responses to building facades employ a correlational strategy involving assessments of actual buildings. Stamps's research strategy involved the use of computer-generated sketches of building facades that systematically varied the architectural treatment of each facade. Based on a previous pilot study, four key variables were identified as having a potential impact on respondent assessments; these variables were: visual area, partitioning of facade elements, fenestration, and articulation (e.g., bays or notches) of the facade plane. Using an experimental research protocol that enabled multiple treatments to be combined across a limited number of stimuli (i.e., the facades), Stamps generated the nine facade examples represented in Figure 9.17. A survey research firm was asked to recruit a random selection of respondents from the area. Each respondent was asked to view paired sets of the facades and indicate which facade appeared to be more massive.

The results of Stamps's study indicate that the most influential variable by far was visual area, which can be modified in situ by setback requirements. Fenestration treatments had a much more modest impact on perception of mass; and both articulation of the facade plane and the partitioning of facade elements had minimal impact.

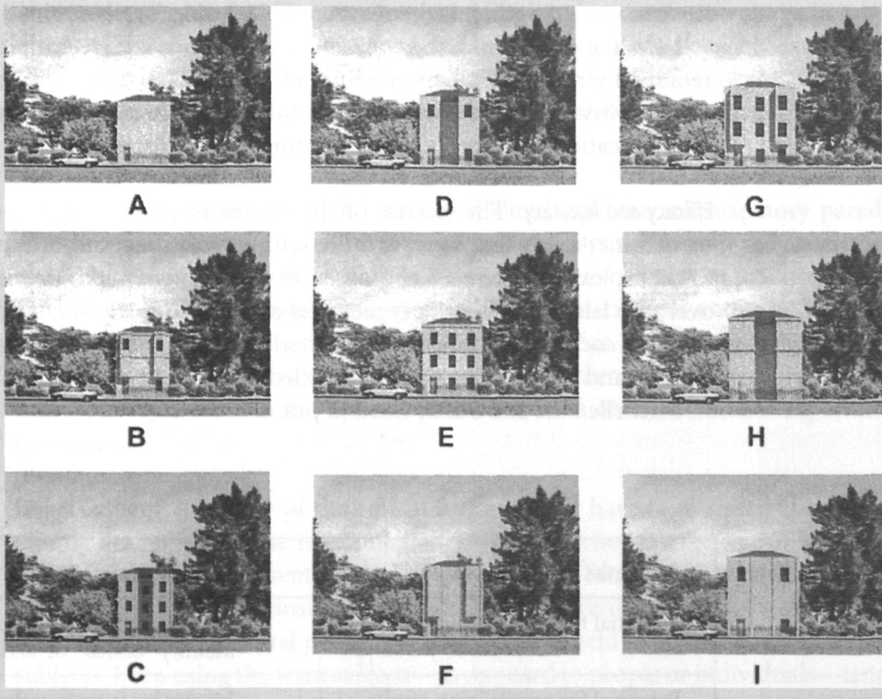

Figure 9.17 Computer-generated facade stimuli from Arthur Stamps. Courtesy of Pion Limited, London.

*Arthur Stamps, "Measures of Architectural Mass: From Vague Impressions to Definite Design Features," *Environment and Planning B: Planning and Design* (1998): 825–836.

9.6 CONCLUSIONS: STRENGTHS AND WEAKNESSES

Of all the research design strategies commonly employed by researchers, the experiment is, in all likelihood, the most controversial. On the one hand, experimental design is considered by postpositivist researchers to represent the highest standard of research.

> The best method—indeed the only fully compelling method—of establishing causation is to conduct a carefully designed experiment in which the effects of possible lurking variables are controlled. To experiment means to actively change {x} and observe the response {y}.[15]

This quotation crisply encapsulates the essence of what is seen as experimentalism's major strength: it is the most credible device for determination of causality.

On the other hand, the experimental design is widely criticized for a variety of reasons by researchers representing both the naturalistic and emancipatory paradigms. Most criticism centers on the following issues: 1) efficacy and accuracy; 2) misapplication of experimental procedure, or 3) ethical concerns. (See Figure 9.18.)

Efficacy and Accuracy. The essence of the argument concerning the efficacy of experimental method is that most real-life settings and socio-cultural phenomena are far too complex to be reduced to a small set of treatment and outcome variables. Moreover, the laboratory setting is seen less as a "neutral social environment" than as a "specific social environment that exerts its own effects."[16] Critics argue that instead, settings and phenomena must be studied in natural settings, in all their complexity. As Michelle Fine and Susan Gordon put it,

Strengths	Weaknesses
Potential for establishing causality	Reduction of complex causality reality to identify "casual" or independent variables
Potential for generalizing results to other settings and phenomenon	Misuse by overgeneralization to different ethnic, gender populations
Ability to control all aspects of experimental design enables attribution of causality	Overemphasis on control yields ethical problems, dehumanization

Figure 9.18 Strengths and weaknesses of experimental research.

> If you really want to know either of us, don't put us in the laboratory, or hand us a survey, or even interview us separately alone in our homes. Watch me (MF) with women friends, my son, his father, my niece or my mother and you will see what feels most authentic to me. These very moments, which construct who I am when I am most me, remain remote from psychological studies of individuals or even groups.[17]

Misapplication. Critics who cite the misuse or misapplication of experimental protocol frequently focus on the way biases or oversights can inadvertently influence the results of such research. This critique is articulated quite clearly by the well-known feminist researcher Shulamit Reinharz. She argues:

> [P]ublication practices and experimental design highlight differences and hide similarities between groups. Overgeneralization that masks differences in race, age, education, and other factors is clearly inappropriate and possibly dangerous. Too often studies done on white populations are generalized to all groups, just as studies done on men are generalized to all people, thereby producing distorted results.[18]

A number of feminists and others affiliated with the emancipatory paradigm have proposed a more nuanced and pragmatic perspective whereby the experimental research design is actually employed to reveal gendered and racist practices. Indeed, Devlin's study of gender discrimination in hiring is an example of this trend. Implicit in this use of the experimental method is the belief that, given the power and respect it commands in so many quarters, feminist and other emancipatory research will only be seen as credible if it is conveyed in the form of the dominant experimental paradigm.

Ethical Concerns. The core of the ethical concerns that have been raised about experimental design is that the manipulative control exercised by the researcher puts research subjects in an essentially powerless position. Treatments are often applied to subjects without their consultation. A potentially advantageous treatment (i.e. better lighting or gender-neutral pedagogy) might be withheld from the "control" group of subjects. Even using the term *subjects*—as opposed to people or individuals—tends to dehumanize those who participate in such studies.

In the end, it would seem that the experimental research design offers both profound strengths and potentially serious weaknesses. The former include the ability to attribute causality, as well as prestige and credibility in some circles. Indeed, in some areas of research—notably in the more technical areas—the premises of experimental work remain unchallenged, although such work is now frequently complemented by computer simulation models. The shortcomings of the experimental model include inappropriate simplification of complex research issues; the potential for misapplication; and the potential for serious ethical problems. However, even Reinharz argues that despite its apparent weaknesses, researchers may do well to exploit its strengths:

> Combining the strengths of the experimental method with the strengths of other methods is probably the best way to avoid its weaknesses while utilizing its power. Similarly, combining the strength of research with the power of other forms of persuasion is probably a useful approach for creating change.[19]

The notion of combining distinctly different research strategies is one that has become increasingly popular among researchers in diverse fields and disciplines. It is a topic to which we will return in Chapter 12.

9.7 RECOMMENDED READING

For readers who would like a general discussion of experimental and quasi-experimental research design, there are two classic books: Donald T. Campbell and J. C. Stanley, *Experimental and Quasi Experimental Designs for Research* (Chicago: Rand McNally, 1966); and Thomas D. Cook and Donald T. Campbell, *Quasi-Experimentation: Design & Analysis Issues for Field Settings* (Boston: Houghton-Mifflin, 1979).

For readers who would prefer a more compressed overview of experimental and quasi-experimental design, Donna Merten's chapter "Experimental and Quasi-Experimental Research," *Research Methods for Education and Psychology* (Thousand Oaks, Calif.: Sage Publications, 1998) is quite readable. A briefer, more prescriptive reference on how to conduct experimental research is provided by John Creswell in a chapter segment entitled "Components of an Experimental Method Plan," in *Research Design: Qualitative and Quantitative Approaches* (Thousand Oaks, Calif.: Sage Publications, 1994): 126–142.

NOTES

1. Baruch Givoni, Michael Gulich, Carlos Gomez, and Antulio Gomuz, "Radiant Cooling by Metal Roofs in Developing Countries," *Proceedings of the 21st National Passive Solar Conference* (Boulder, Colo.: American Solar Energy Society, April 1996): 83–87.
2. Ann S. Devlin, "Architects: Gender-Role and Hiring Decisions," *Psychological Report* 81 (1997): 667–676.
3. Ibid., 670.

4. Ibid., 667.
5. Ibid., 674.
6. Thomas D. Cook and Donald T. Campbell, *Quasi-Experimentation: Design and Analysis* (Boston: Houghton Mifflin, 1979).
7. Ibid.
8. Janice Barnes, Kaninika Bhatnagar, Fernando Lara, Satoshi Nakamura, Pirasri Povatong, Tien Chien Tsao, and Victoria Turkel, "Results of a Quasi-Experimental Treatment in the Architecture Gallery" (unpublished student paper, University of Michigan, 1997).
9. Despite these conclusions, a major obstacle to changing the furniture arrangement on a permanent basis was that the intended lack of visibility meant the security of the exhibits could not be monitored from the offices across the hall. As a result of this quasi-experiment, then, the student group made a policy recommendation to the administration that a separate student/faculty lounge area should be provided. At this writing, more than three years later, a new lounge is being built.
10. John Creswell, *Research Design: Qualitative and Quantitative Approaches* (Thousand Oaks, Calif.: Sage Publications, 1994); D. Mertens, *Research Methods in Education and Psychology* (Thousand Oaks, Calif.: Sage Publications, 1998).
11. Edward Arens, Tengfang Xu, Katsuhiro Miura, Zhang Hui, Marc Fountain, and Fred Bauman, "A Study of Occupant Cooling by Personally Controlled Air Movement," *Energy and Building* 27 (1998): 45–59.
12. Ibid., 46–47.
13. Ibid., 47.
14. Saso Medved and Peter Novak, "Heat Transfer through a Double-Pane Window with Insulation Screen Open at the Top," *Energy and Building* 28 (1998): 257–268.
15. D. Moore and D. McCabe, *Introduction to the Practice of Statistics* (New York: Freeman, 1993).
16. Shulamit Reinharz, *Feminist Methods in Social Research* (New York: Oxford University Press, 1992), 100.
17. Michelle Fine and Susan M. Gordon, "Feminist Transformations of/Despite Psychology," in *Gender and Thought: Psychological Perspectives*, ed. M. Crawford and M. Gentry (New York: Springer-Verlag, 1989), 106
18. Reinharz, 107.
19. Ibid., 108.