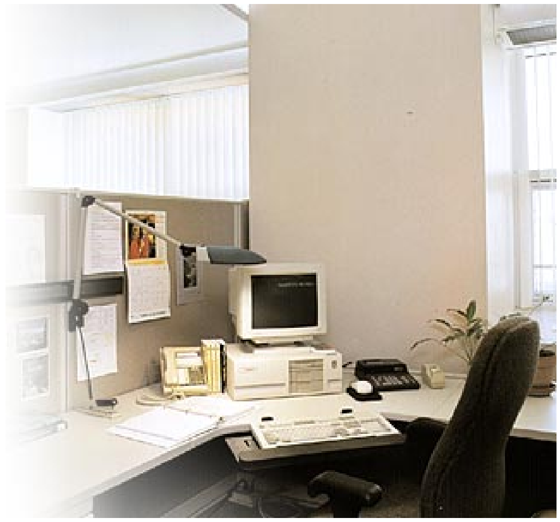

Reviews of
Technical Reports on Daylight and Productivity

Reviews by:

Peter Boyce, Ph.D.
Lighting Research Center
Rensselaer Polytechnic Institute
21 Union St.
Troy, NY 12180

For: The Daylight Dividends program

The Daylight Dividends program (www.daylightdividends.org) was established to build market demand for daylighting as a means of improving indoor environmental quality; to overcome technological barriers to effectively reap the energy savings of daylight; and to inform and assist state and regional market transformation and resource acquisition program implementation efforts.

Daylight Dividends Sponsors

California Energy Commission
Connecticut Light & Power
Iowa Energy Center
Lighting Research Center
North Carolina Daylighting Consortium
Northwest Energy Efficiency Alliance
NYSERDA
US Department of Energy

NOTE:
This report was prepared by the Lighting Research Center with support of the U.S. Department of Energy, under Award No. DE-FC26-02NT41497 and from the California Energy Commission, The Connecticut Light & Power Company, New York State Energy Research & Development Authority, Northwest Energy Efficiency Alliance and the North Carolina Daylight Consortium. However, any opinions, findings, conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of DOE or other aforementioned parties.

Lighting Research Center

Contents

Technical Reports on Daylight and Productivity: Some Observations

Three reviews by Peter Boyce, Ph.D., Lighting Research Center, May 2004

In October 2003, the California Energy Commission (CEC) released three new reports on the topic of productivity and interior environments. These reports were prepared by the Heschong-Mahone Group under the Public Interest Energy Research (PIER) program. These reports describe epidemiological studies of the impact of daylight on: (1) the performance of office workers, (2) merchandise sales in a retail store chain, and (3) the progress of elementary schoolchildren. All three reports use similar methodology for seeking a statistical relationship between numerous explanatory variables and a desirable outcome. Data are collected from databases or on-site surveys and multiple linear regression equations are developed between the explanatory variables and the desired outcomes. The technical reports concerning retail sales and learning in schools attempt to replicate previous findings. These attempts failed.

Reviews of the individual reports follow this summary. However, a number of observations can be made about the three reports as a set, covering both methodology and the effects of daylight. These observations include:

- Attempts at replication are valuable. Effects that can be replicated can be relied on. Effects that cannot be replicated cannot be relied on.
- Variables in multiple linear regression equations that explain very small amounts of variance in the data are not to be trusted; they are inherently unstable. They are likely to change their effect in both magnitude and direction if different variables and different datasets are used. Any conclusions reached about such variables are built upon sand.
- Epidemiology can only establish a correlation, not a cause. To understand what causes an effect it is necessary to understand the mechanisms by which an independent variable, in this case, daylight, might effect the dependent variable, in this case either office worker performance, or retail sales, or school children's' progress in math or reading.
- Attempts to understand what these mechanisms might be show that simply providing daylight is not a guarantee of success. Consistent with the view expressed in the Daylight Dividend review entitled "The Benefits of Daylight through Windows", whether daylight has a positive or negative effect on the outcomes studied depends on how it is delivered. Daylighting that provides an even distribution of daylight, reveals an extensive view, has controls to limit glare and thermal heat gain, and does not introduce distraction, is likely to make a positive contribution. Daylighting that causes visual or thermal discomfort, contributes to auditory difficulties, or introduces distraction, is likely to make a negative contribution.
- You do not need to be a "rocket scientist" to understand these aspects of windows and skylights. These aspects are well understood. Given knowledge of the specific situation in which the windows or skylights are to be installed, they can be handled by any competent designer.

Reviews of each report, with appropriate references, follow.

Review of Technical Report: Windows and Offices: *A Study of Office Worker Performance and the Indoor Environment*

Review by Peter Boyce, Ph.D., Lighting Research Center, May 2004

The technical report, *Windows and Offices: A Study of Office Worker Performance and the Indoor Environment*, was prepared by Heschong-Mahone Group in October 2003. The report describes two field studies undertaken to determine how indoor environmental conditions might influence office worker performance. Both studies are epidemiological, seeking a statistical relationship between numerous uncontrolled variables and measures of office worker performance. Both use a similar statistical approach to that adopted in earlier studies by this group on the effect of daylight on retail sales[1] and on school students' test scores[2].

The first study was set in the Sacramento Municipal Utility District (SMUD) call center and examined the impact of many variables on the average handling time for incoming calls for 100 office workers. The variables covered lighting, view, temperature and ventilation, call center status (number of calls answered, etc.), personal status, and work group. Data were collected for four weeks in September 2002 and three weeks in November 2002. Multiple linear regressions, using a variant of the backward elimination method, were fitted for daily data in September and November and for hourly data in November only. These three regression equations form three models of the relationship between the variables measured and the average handling time for incoming calls.

Examination of these models shows that a number of variables are inconsistent in that they have statistically significant effects in one model but not in another. However, there are seven variables that have a statistically significant effect on average handling time that is greater than 1% and in the same direction for all three models. The biggest effect predicted by the regression is a 17% to 19% reduction in average handling time. This has to do with whether an individual is a team leader. Team leaders show shorter handling times than other workers, which presumably is one reason why they are team leaders. The next biggest predicted effect is cubicle partition height; workers with the highest partitions tend to have 11% to 18% longer average handling times than those with the lowest partitions. This detrimental effect may be due to the partition height restricting view, but this is open to question because the team leaders have low partition heights. This means that partition height may be confounded with team leaders and the effect may have nothing to do with view. The next three biggest predicted effects are shown by three distinct work groups (3% to 17%; 4% to 14%; 5% to 15%), which simply tells us that some work groups in the call center are faster than others. The results then get more interesting in that the next biggest predicted effect is that of view when looking around from the cubicle (6% to 7%); workers who can see out through a large, unobstructed window tend to have shorter handling times than those with no view out. The final variable that shows consistent effects is the position of the register of the floor air supply (6%): those who have the register open to provide a supply of conditioned air tend to have shorter average handling times than those who have the register closed (3% to 10%). Among the variables that have no consistent, statistically significant effects on average handling time are the levels of daylight or electric light and the outside view available when the worker is looking at the monitor.

It is important to appreciate that these effects are predicted from the regression equations by holding all but one of the variables in the equation constant, and then changing the value of that variable over its maximum range, e.g., from no view to a view out that fills the worker's visual field. To understand how reliable such predictions are likely to be, it is necessary to consider how good the regression equation fits to the data. Goodness of fit can be quantified by the amount of variance in the data that can be explained by the regression equation. For the September daily data, 21% of the variance is explained by the regression equation; for the November daily data, 22% of the variance is explained; for the November hourly data only 8% of the variance is explained by the regression equation. Of course, these percentages are for the complete equation. The amount of variance explained by introducing partition height is less than 1% for all three data sets. Similarly, the amount of variance explained by introducing view is less than 0.7% for all three data sets. In terms of Cohen's protocol for small, medium, and large effect sizes[3], these effects of partition height and view are less than small. This means that the effects of these variables are likely to be difficult to replicate.

The second study administered five cognitive tests via the SMUD computer system to 201 people working in three different SMUD buildings. The buildings offer a wide range of access to daylight and view. The tests ranged from those measuring visual capabilities, such as visual acuity and visual search, to those measuring short-term and long-term memory. The tests were administered five times—in the middle of the day on successive Thursdays from mid-October to November. Again, an epidemiological approach was used to develop multiple regression equations to relate many variables to the performance on the different tasks. For the tests related to visual capabilities, the effects were, as would be expected from what is known about vision and visibility (i.e., the time taken to identify a visual acuity target and the time taken to find the target in a visual search task), that both increase with increasing age (13% and 6%, respectively) and decrease with the use of a high-resolution monitor (29% and 16% respectively). The effect of having a wider view out or of sitting under a skylight is to increase the time taken to identify the orientation of a visual acuity target (13%) and the time taken to find the target in a visual search task (15%), probably because of reduced contrast due to the higher levels of ambient light being reflected from the monitor. Interestingly, having a good view out tends to improve performance on the memory tasks (10% to 16%) while glare from windows tends to reduce it (15% to 17%). This latter effect might be explained by the glare reducing the contrast of the stimuli used for the memory tasks. These results serve to emphasize the fact that the impacts of windows can be both good and bad, and can operate directly by changing visual conditions and emotionally by changing mood.

These effects come from the regression equation for each task, so it is necessary to know how good the fit of the regression equation is to the data. The answer for these tasks is poor, with one exception. For the visual acuity, letter search, number search, and short-term memory task, the percentage of variance explained by the complete regression equation ranges from 11% to 20%. For the long-term memory test, the variance explained is 51% but of this 42.5% comes from two variables related to the correctness of the responses shortly after seeing the articles to be memorized, i.e., short-term memory. The variances explained by the addition of the view and glare variables are less than 1% for the regression equations for all tasks, except for one which rises to 1.4%. Again, according to Cohen's protocol for small, medium, and large effect sizes[3], the effect sizes of the view and glare variables are less than small.

Finally, a questionnaire was administered to the subjects in the second study. This revealed that people who had better views also tended to be more positive about their health, and reports of increased fatigue were associated with a lack of view, although the correlation coefficients between these variables were all low (typically about 0.2). Interestingly, the content of the view (more sky, more trees, or human activity) was equally as important for these relationships as the size of the view.

Overall, this is an interesting, clearly written report that shows both pluses and minuses for windows as a factor in determining the performance of office workers. It also gives full details of some of the difficulties experienced in measuring the physical environment, it discusses some of the alternative explanations for the results, and it mentions some of the aspects of the physical environment that might be influential but were not included. No doubt the executive summary will be widely read and quoted, but one needs to be careful to understand some of its wording. For example, it quotes some remarkable percentage increases in office worker performance achieved by providing a good view without making clear that these are predictions based on a statistical model and are the result of combining the effects of two view variables. It also claims that the physical environment explains about 2% to 5% of the total variance. This is correct, but the physical environment includes aspects of the environment other than windows and daylight. If only the variables related to daylight and windows are included, then the variance explained is only 0.2%

to 1.4% for the call center study and from 0.4% to 4.5% for the cognitive tests study. Similarly, the executive summary states that the physical environment represents about 1/8 to 1/3 of our entire ability to predict variation in individual worker performance. No doubt this is true for this study, but it could presumably be increased to 100% if only variables connected to the physical environment had been included in the model, although then the percentage variance explained by the equation as a whole would presumably be even less. To fully appreciate all the intricacies of this report, together with its implications for building design and future research, it is necessary to read the full report.

References

1. Heschong-Mahone Group (1999) "Skylighting and retail sales: An investigation into the relationship between daylight and human performance. Detailed Report." Fair Oaks, CA.

2. Heschong-Mahone Group (1999) "Daylighting in schools: An investigation into the relationship between daylight and human performance. Detailed Report." Fair Oaks, CA.

3. Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences. Hillside, NJ: Lawrence Erlbaum Associates.

Review of Technical Report:
Daylight and Retail Sales

Review by Peter Boyce, Ph.D., Lighting Research Center, May 2004

The technical report, *Daylight and Retail Sales*, was prepared by Heschong-Mahone Group in October 2003. The report is a follow-up from an earlier study concerned with the influence of skylights on sales in a chain of retail stores.[1, 2] In this earlier study, sales data were obtained from 108 stores in a chain of stores selling similar goods and laid out in a similar way. Two-thirds of the stores had diffusing skylights that provided daylight, one-third did not. A multiple linear regression analysis was undertaken of the relationship between sales in the stores and twelve explanatory variables including the presence or absence of skylights, the size of the store, the hours the store was open, the size and income of the surrounding population, and the number of years since the store was remodeled. The resulting regression equation explained 58% of the variance in sales. The most important variable for predicting sales was the hours the store was open per week; this variable explained 16% of the variance in sales. The next most important variable was the number of years since the last remodeling, which explained 9% of the variance. The presence or absence of skylights was the third most important variable for predicting sales, explaining 4% of the variance. In terms of Cohen's protocol of effect sizes which is based on percentage variance explained,[3] the presence or absence of skylights variable has a small effect. However, in terms of the magnitude of the predicted effect on sales, the effect is much more impressive. Specifically, the regression equation predicted that adding skylighting to the average non-skylit store in this chain would increase sales in the store by 40%. Such a dramatic increase in sales would be irresistible to a retailer and would lead to much more widespread use of skylights in stores. This combination of a small effect size, based on variance explained, and a large effect with important financial implications, based on the predictions from the regression equation, supports the need for a further study to replicate these results. This technical report describes such a study.

In this study, a similar but more sophisticated evaluation was carried out on a different chain of stores operating in a different retail sector in California. In total, sales from 73 stores were examined, 24 of them having a significant amount of daylighting, primarily from diffusing skylights. Sales data were provided by the retailer for 34 months during the years 1999 to 2001. This data was separated into two groups, 24 months for 1999 and 2000, and 10 months during 2001. This division was made because during 2001, California was suffering from power shortages, so most stores were operating their electric lighting at about half power. This meant that during the ten months of 2001 there would have been a more obvious difference between the daylit and non-daylight stores than during the 24 months of 1999 and 2000, during which the electric lighting in the non-daylit stores was run at full power and the electric lighting in the daylit stores was dimmed when there was plenty of daylight.

For these two time periods, the study sought to establish statistical relationships between two dependent variables and a total of fifty independent variables. The two dependent variables reflect the desired outcome, one being an index of monthly sales per store and the other being an index of monthly transactions per store. The independent variables reflect the possible explanations for any changes in sales and transactions. The independent variables can be divided into five groups: corporate data relating to the size, age, and hours of operation of the store; census data relating to the population size, growth, and ethnic make-up in the surrounding area; local market influences relating to the extent of competition, visibility of the store, parking facilities, etc.; store comfort conditions relating to the amount of daylight, vertical illuminance distributions, ceiling height, noticeable air movement, smell, noise, and cleanliness of the store; and interactions between some of these variables and the amount of daylight.

Given the large number of independent variables, the first step in the statistical analysis was to identify which of the corporate, census, and market variables had a statistically significant effect ($p<0.10$) on the sales index. This approach showed that for both the 24-month and 10-month sales data, the statistically significant variables were: the size of the store; its hours of opening and age; the population density, population growth, and the household status in the surrounding area; the presence of sister stores or competitors; and the amount of parking.

Having established these core variables, the next step was to replicate the previous study by introducing the presence or absence of daylighting as an independent variable.[1, 2] Despite the fact that the previous study had predicted a major effect on sales of providing daylight in the store, the presence or absence of daylight did not have a statistically significant effect on the sales index for either the 24-month data or the 10-month data. The variables that had a statistically significant effect on sales in both 24-month and 10-month models were largely what would be expected: store size, age, and hours of opening; the presence of sister stores and competitors in the area; and the availability of parking.

Having failed to replicate the previous study, the next step was to devise a more precise way to quantify the daylight conditions rather than simply as a presence or absence. The measure chosen was the number of hours per year that the illuminance provided by daylight would exceed the illuminance provided by electric lighting in the non-daylit stores. This measure gives a single number that reflects both the duration and intensity of daylight. It was calculated for each store using the details of the daylighting system used, the store interior reflectances, and typical meteorological year data. In addition, interaction terms between daylight and the other significant variables listed above were introduced. Finally, both linear and natural logarithmic models were developed. Logarithmic models are appropriate where the effect of adding daylight is expected to show a law of diminishing returns. Linear models are appropriate when it is not.

Both linear and logarithmic models explain a high percentage of the variance in the sales data (see Table 1), with little difference between them. A large part of this variance is explained by the age of the store in both models (28% to 38% variance explained). Surprisingly, these models predict that the older the store is the more it sells, although whether this is due to a correlation between store size and age or whether it is simply that the older the store the more likely it is to have been remodeled is not clear.

Model Type and Outcome Variable	24-Month Data	10-Month Data
Linear / Sales Index	75.3%	76.5%
Log / Sales Index	75.7%	74.7%
Linear / Transaction Index	75.2%	77.2%

Table 1: Percentage of the total variance in sales index data and in number of transactions index data explained by the linear and logarithmic models (only linear models were developed for the transactions index data)

Without further explanation, this finding conflicts with the result of the earlier study in which it was found that the larger the number of years since the last remodeling the greater the reduction in sales index. The size of the store and the availability of parking are the next most important variables in the two models (5% to 8% variance explained and 4% to 8% variance explained, respectively). The larger the size of the store and the greater the amount of parking, the higher is the sales index. The number of daylight hours was statistically significant in both linear and logarithmic models, for both datasets and for both sales and transactions indices. The amounts of variance explained by the number of daylight hours and by the statistically significant interactions involving daylight are given in Table 2.

Model Type / Outcome Variable / Independent Variable	24-Month Data	10-Month Data
Linear / Sales Index / Daylight Hours	5.0%	5.3%
Log / Sales Index / Daylight Hours	3.9%	3.8%
Linear / Sales Index / Daylight Hours X Age Interaction	-	1.7%
Log / Sales Index / Daylight Hours X Age Interaction	-	1.6%
Linear / Sales Index / Daylight Hours X Parking Interaction	0.1%	0.8%
Log / Sales Index / Daylight Hours X Parking Interaction	0.4%	1.3%

Table 2: Percentage of the total variance in sales data explained by the daylight hours variable and the interaction variables involving daylight (the daylight hours age interaction was not statistically significant for the 24 month data)

These variance-explained values for daylight hours are similar to that obtained in the previous study for the presence and absence of daylight.[1,2] In terms of Cohen's protocol of effect sizes, these are small effects. However, it is important to note that in terms of variance explained, the daylight hours variable is comparable in size with such variables as the population density of the surrounding area and the number of nearby competitors, both variables that are considered important in real estate analysis.

For the predictions made by these models, Table 3 lists the mean chain-wide percentage increases in sales index and transactions index predicted by the different models for the introduction of daylight into the stores.

Model Type and Outcome Variable	24-Month Data	10-Month Data
Linear / Sales Index	-0.3% (-50% to +56%)	+5.2% (-46% to +88%)
Log / Sales Index	+1.1% (-45% to +37%)	+5.7% (-35% to +49)
Linear / Transaction Index	+1.2%	+2.1%

Table 3: Predicted mean chain-wide percentage changes in sales index (and the associated range) and in number of transactions index from the linear and logarithmic models (only linear models were developed for the transaction index data.)

These predicted mean percentage changes are a long way from the 40% increase in sales index predicted in the earlier study.[1,2] In fact, the means are so close to zero and the ranges so large that the report admits it is possible that all of the mean percentages could change sign if a different population of stores had been used. Further, the coefficient for the daylight hours variable is negative in all of the models. This is interesting because it implies that the greater the number of daylight hours, the greater the reductions in sales index and number of transactions index. It is the interaction with the availability of parking that makes the net effect of providing some daylight positive. Specifically, for the log model applied to the 10-month data, which is reported in detail, it was found that when the availability of parking is one standard deviation greater than normal, the mean effect of adding daylight is enhanced to +19.7%, and when the availability of parking is one standard deviation below the norm, the mean effect of adding daylight reduces the sales index by -8.7%. This suggests that adding daylight to a store will only have a positive effect when there is sufficient parking to deal with the additional demand.

One advantage of expressing the availability of daylight as the number of hours of daylight above a threshold illuminance is that it makes it possible to test for a dose/response relationship, (i.e., does more daylight mean more sales?). Examining the predictions of the logarithmic models while holding the level of parking constant always shows increases in sales index as more daylight hours are added. The magnitude of the increase is dependent on the level of parking assumed, as shown in Table 4.

Parking	24-Month Data	10-Month Data
Fixed at mean	+0.4% to +2.0%	-2% to +14%
Fixed at one standard deviation above norm	+5.0% to +27%	+2% to +37%

Table 4: Range of mean chain-wide percentage increases in sales index predicted by the logarithmic models, for different levels of parking, for increases in daylight hours from 250 hours to 1900 hours

From an examination of Tables 3 and 4 it is interesting to note that the predicted effects of adding daylight are almost always greater for the 10-month data than for the 24-month data. As explained above, during 2001 when the 10-month data was collected, the electric lighting in virtually all stores was operated at reduced power because of power shortages. Such an event would make the difference between the daylight stores and the non-daylight stores more obvious and presumably enhance the effect of daylight provision on sales index. However, it should be appreciated that the 24-month data are representative of normal operation.

Finally, one curious result deserves mention. All of the above models used the average of the monthly sales index, but the amount and duration of daylight also varies with the season. Given the demonstrated effect of daylight hours on sales index, it would be reasonable to expect a difference between sales in the months of July and January, both the intensity and duration of daylight being less in January than in July. Unfortunately, for simplicity no difference was found in the models constructed using the January and the July monthly sales index data. The authors suggest that this absence of a seasonal effect occurs because providing daylight in a store inspires long-term customer loyalty, thereby attracting sales throughout the year. A corollary to this suggestion, necessary for the effect of increasing daylight hours to hold, is that more daylight hours inspire loyalty in a greater number of people.

In addition to modeling the relationships between a wide range of independent variables and a sales index, and a number of transactions index, surveys were administered to employees and interviews were conducted with store managers. The surveys revealed that the employees considered the daylit stores to have slightly better lighting quality than the non-daylit stores in terms of visual comfort, brightness, uniformity, and visibility of detail and color. The managers, during 1999 and 2000, perceived little difference in the level of lighting quality of the daylit and non-daylit stores. However, in 2001, when the electric lighting was being operated at half power, many more managers thought the lighting in the daylit stores was good relative to the number who thought the lighting of the non-daylit stores was good. The managers were also asked about their experience with the skylights. Of the 24 managers of daylit stores, six made strongly positive comments about the skylights, while one made negative comments. As for difficulties experienced, four reported that a skylight had leaked once, and one reported that a skylight had been broken for one day. None reported any problems with break-ins or accidental falls.

Finally, the energy savings associated with the use of daylight in the stores were estimated. The lighting energy savings from skylights and photo-control operation of the electric lighting were calculated to be in the 20%-30% range and in the 15%-25% range for lighting and HVAC savings together.

Overall, this is an interesting, clearly written report that repays careful reading. It contains two main messages. The first is that providing daylight in stores operating in the retail sector examined may, on average, have a small but positive effect on sales and number of transactions. The second is that to achieve such gains in sales and transactions it is not enough simply to install skylights. Other factors must be in place. For example, the parking at the store has to be sufficient to accommodate the increase in the number of customers. If it is not, then providing daylight may lead to a reduction in sales.

References

4. Heschong-Mahone Group (1999). "*Skylighting and retail sales: An investigation into the relationship between daylight and human performance. Detailed report*." Fair Oaks, CA.

5. Heschong, L., Wright, R.W., and Okura, S., (2002). Daylighting impacts on retail sales performance, *Journal of the Illuminating Engineering Society*. 31, 21-25.

6. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillside, NJ: Lawrence Erlbaum Associates.

Review of Technical Report:
Windows and Classrooms: A Study of Student Performance and the Indoor Environment

Review by Peter Boyce, Ph.D., Lighting Research Center, May 2004

The technical report, *Windows and Classrooms: A Study of Student Performance and the Indoor Environment*, was prepared by Heschong-Mahone Group in October 2003. The report describes a continuation of two earlier studies concerned with the influence of daylight on the performance of students.[1,2,3] In the first of these studies the impact of having daylight in classrooms on the performance of elementary school children on standardized tests was examined for three school districts.[1,2] The school districts were in three different states each having different climates, different building types, different curriculums, and different testing protocols. In total, the performances of about 8,000 students were examined in each district. For two of the districts, Seattle, Washington, and Fort Collins, Colorado, a multiple linear regression analysis was undertaken of the relationship between end-of-year performances on math and reading tests and multiple explanatory variables. One of those was a daylight code reflecting the combined effects of windows and skylights on daylight provision. These regression equations showed a statistically significant effect of daylight ($p<0.01$), the students in classrooms with the most daylighting having a predicted 7% to 13% higher test scores than those whose classrooms had the least daylight. In the third district, Capistrano, located on the coastal plain of California, it was possible to examine the progress students made over the year because this district administered tests in the fall and spring. For this district, a multiple linear regression analysis was undertaken of the relationship between the difference in performance on standardized math and reading test scores between fall and spring and fifty explanatory variables, one of which was a daylight code that reflected the combined effects of windows and skylights on the daylight provision. These regression equations also showed statistically significant effects of daylight ($p<0.01$), the students in classrooms with the most daylighting progressing 20% faster in math and 26% faster in reading than those whose classrooms had the least amount of daylight.

These remarkable results generated a lot of interest and controversy, so much so that a re-analysis of the Capistrano data was undertaken in an attempt to rule out some of the variables that might have been confounded with the amount of daylight, and hence might be the real reason for the apparent effect of daylight.[3] After looking at the possibility that better teachers were assigned to classrooms with more daylight, that absenteeism and tardiness were linked to the amount of daylight, that teachers had a strong preference for daylit classrooms, and that different grades would show different effects of daylight, it was concluded that none of these variables could explain the impact of daylight on student performance that had been found. However, it remains a fact that the regression equations describing the impact of daylight explain only 26% of the variance in the data, and the daylight variable itself explains only 0.3% of the variance. In terms of the Cohen's protocol for assessing effect sizes expressed as percentage of variance explained,[4] this effect of daylight is less than small. Nonetheless, this combination of a small effect size, based on variance explained, and a large effect with important educational and societal implications, based on the predictions from the regression equation, supports the need for a further study to replicate these results. As the authors state, "For any scientific study, regardless of the strength of the initial findings, replication is the acid test of validity." This technical report describes such a study, undertaken with the objective of answering the question, "Will we find similar results, that more daylight in classrooms is associated with faster student learning, using the same methodology, but with a different study population?"

In this study a similar but more sophisticated study was undertaken in the Fresno school district, located in the central valley of California. This large school district tests students in both fall and spring, has an extensive database of demographic information, has a wide variety of classroom daylighting conditions, and has different climate and architectural conditions than those of the Capistrano school district. In total, the changes in performance on standardized reading and math tests over the school year were collected for approximately 8,500 elementary students. Further, data on 150 explanatory variables were collected either from the district's databases or during visits to the 500 classrooms occupied by the classes participating. These 150 variables can be grouped into those concerning the school site, e.g., age, student population, neighborhood; student characteristics, e.g., enrolled in gifted and talented or special education program, English learners, grade level, ethnicity and gender; teacher characteristics, e.g., years teaching in district, salary level, credential status, gender, and ethnic status; school demographics, e.g., student mobility, parent education, free lunch provision; window and daylight characteristics, e.g., daylight code, window area, window view; classroom characteristics, e.g., size, type, amenities; indoor air quality, e.g., HVAC system type, HVAC controls, smell; noise, e.g., HVAC noise, percentage acoustic wall; and electric lighting, e.g., luminaire type and condition, illuminance, and lamp color.

From these data, the study sought to establish statistical relationships between two dependent variables and the 150 independent variables. The two dependent variables reflect the desired outcome, one being progress in math, the other progress in reading. The independent variables reflect the factors that might be expected to influence learning and hence progress in math and reading. Given the large number of independent variables, the first step in the statistical analysis was to identify which student and teacher characteristics had a statistically significant effect ($p<0.10$) on the students' progress. Ten student characteristics and five teacher characteristics were consistently statistically significant in relating to the progress in math and reading and were therefore included in a base demographic model. The ten student variables included were the grade level; the percentage attendance; whether the student was qualified or enrolled in a gifted and talented education program; whether the student was a special education student; the level of English language development; whether the student was the recipient of a free lunch (there is such a thing as a free lunch); whether the student was the recipient of a reduced cost lunch; whether the student was living in a non-standard situation, e.g. foster care; the student's gender; and the student's ethnicity. The five teacher variables were the teacher's annual salary; the number of years the teacher had been employed in the Fresno Unified School District; whether the teacher was a mentor teacher; whether the teacher was a pre-tenure teacher; and whether the teacher was responsible for more than one grade level in the same classroom.

Having established these base variables, the next step was to replicate the previous study[1,2] by adding six variables similar to those that had been found to be statistically significant in the Capistrano study (operable windows, portable classroom, open classroom, classroom size, school population, school age) together with the daylight code variable that quantified the daylight provision. Despite the fact that the Capistrano study had predicted a major effect of daylight on progress in math and reading, the daylight code did not have a statistically significant effect on progress in math and reading in Fresno.

Having failed to replicate the previous study the next step was to retreat to the base demographic variables and then add eight different groups of independent variables, one by one, with and without the daylight code. The eight different groups of variables were based around the window characteristics, the air quality, the noise, the electric lighting, the classroom characteristics, the specific school, the school site, and the school socioeconomic status. Adding these different groups of variables had little effect on the variance explained by the basic demographic variables. The specific school group added 1.1% and 1.5% to the variance explained in the reading and math scores, respectively. The classroom characteristics group added 0.3% and 1.1% to the variance explained in the reading and math score respectively. The window group added 0.7% and 0.6% to the variance explained in the reading and math scores, respectively. All of the other variable groups added even less to the variance explained.

Examination of the effects of the individual variables in these groups enabled the authors to select a total of 72 independent variables related to the physical environment to be added to the basic demographic variables, as well as the daylight code, in the final models. Twenty-one variables describing the physical conditions were statistically significant in the final math model, of which seven were window characteristics. The corresponding figures for the reading model were twenty -seven and eight. Of course, statistical significance simply tells you whether the effect of that variable is due to chance or not. It does not tell you anything about the size of the effect. To know that, you have to look at the variance explained. The final math model explained only 19.2% of the variance, and the final reading model explained 25.5%, meaning that 80.8% and 74.5% of the variances in math and reading scores respectively are unexplained by these models. Table 1 shows the variables that explain more than 1% of the variance, in both the final reading and the final math models.

Variable	Math – percent variance explained and direction of effect	Reading – percent variance explained and direction of effect
Fall score	4.3% - negative	1.8% - negative
Enrolled in gifted and talented education	2.8% - positive	-
Fourth grade	1.5% - negative	-
Re-test for fall test	1.2% - positive	-
School English learner	1.0% - positive	1.1% - negative
Special education student	1.0% - negative	-

Table 1: Variables that explain more than one percent of the variance in the final math and reading models, the percentage variance explained and the direction of the effect

The most important single variable in predicting progress in math and reading is the individual student's fall test score; the higher the fall test score, the less is the progress in math and reading, probably because there is less room for improvement. Other important variables have to do with the student's status. All of the physical variables in both models explain less than 1% of the variance, often less than 0.1%. The problem with variables that explain so little variance is that they are very unstable. They can be statistically significant or not, depending on what other variables are included. Even when they are statistically significant, they can change the direction of their effect. For example, the daylight code explained 0.3% of the variance in the Capistrano math model, and had a positive effect, i.e., the more daylight, the greater the progress. In the Fresno math model, the daylight code explains less than 0.1% of the variance and has a negative effect, i.e., the more daylight, the less the progress. In the Fresno reading model, the daylight code explained 0.1% of the variance and had a negative effect, i.e., the less the daylight, the greater the progress in reading.

These unexpected effects of daylight code caused the authors to look closely at how the effects of all the physical variables, related to windows, changed with different daylight codes. For both the math and reading models, it was found that both a high daylight code and a low daylight code were most beneficial for progress in reading and math. Specifically, classrooms with good uniform daylight, a nice view, and controls on the window so that glare was avoided were good for progress, as were classrooms with very little daylight and good control on what windows there were. The worst classrooms for progress were classrooms with some daylight, without a nice view, and with windows oriented so as to experience glare from the sun and without any control on the windows.

These findings inspired the authors to return to a sample of the classrooms to investigate what aspects of the daylit classrooms were apparently affecting student performance negatively. As a result of these investigations it was found that the daylight code was a reasonable metric for the daylighting of the classrooms in use. There was no evidence that there were pervasive lighting quality problems in the daylit classrooms. Similarly, there was little evidence that the thermal comfort and air quality in the classrooms was linked to daylight. Rather, any thermal comfort and air quality problems were associated with the operation of the HVAC system. Where there was some link between adverse environmental conditions and daylight was in the acoustic conditions. Classrooms with extensive windows had longer reverberation times, tended to receive more noise from outside because the windows were sometimes open, and, because of the structure of the school, tended to have some in-class tutoring going on in the classroom at the same time as the lesson. The classrooms with little daylight were arranged so that any special tutoring could take place outside the classroom.

Overall, this is a comprehensive examination of the effect of daylighting in schools. It shows that providing daylight in classrooms is not a magic bullet for enhancing student progress. Whether daylight has a positive or negative effect on student progress depends on how it is delivered. Daylighting that provides an even distribution of daylight, an extensive view, and limits glare and thermal heat gain, is likely to make a positive contribution to student progress. If having more windows increases the noise level in the classroom and hence makes it more difficult to hear the teacher, or if glare from the sun makes it more difficult to see the teachers working, or if what is happening outside is distracting and visible through the window, then more daylight provided through windows is likely to slow students' progress.

References

7. Heschong-Mahone Group (1999). "*Daylighting in Schools: An Investigation into the Relationship between Daylighting and Human Performance. Detailed Report."* Fair Oaks, CA: Heschong-Mahone Group.

8. Heschong, L., Wright, R.W., and Okura, S., (2002). Daylighting impacts on human performance in school, *Journal of the Illuminating Engineering Society*, 31, 101-114.

9. Heschong-Mahone Group (2001). *"Re-Analysis Report: Daylighting in Schools."* Fair Oaks, CA: Heschong-Mahone Group.

10. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillside, NJ: Lawrence Erlbaum Associates.