LSAC RESEARCH REPORT SERIES

Predictive Validity of the LSAT: A National Summary of the 2001–2002 Correlation Studies

Lisa Anthony Stilwell
Susan P. Dalessandro
Lynda M. Reese

Law School Admission Council
LSAT Technical Report 03-01
December 2005

A Publication of the Law School Admission Council

The Law School Admission Council (LSAC) is a nonprofit corporation whose members are more than 200 law schools in the United States and Canada. It was founded in 1947 to coordinate, facilitate, and enhance the law school admission process. The organization also provides programs and services related to legal education. All law schools approved by the American Bar Association (ABA) are LSAC members. Canadian law schools recognized by a provincial or territorial law society or government agency are also included in the voting membership of the Council.

This study is published and distributed by the Law School Admission Council (LSAC). The opinions and conclusions contained in these reports are those of the authors and do not necessarily reflect the position or policy of the Law School Admission Council.

Table of Contents

Executive Summary 1

Introduction 1
The Criterion Variable 2
The Predictor Variables 2
Estimating Validity 2

Results 4
Summary Data 4
Correlation Coefficients 5
Factors Influencing the Magnitude of the Correlation Coefficients 8
Cross Validation Studies 10
Accounting for Variance 11
Trends Over Time 11

Conclusion 12

References 12

Appendix A 14

Executive Summary

Since the Law School Admission Test (LSAT) was first administered, the sponsors of the test have carried out predictive validity studies to evaluate the effectiveness of the test as well as other predictors in determining first-year law school performance. Over the years, increasingly large numbers of law schools have participated in these studies, commonly called correlation studies.

This report presents a summary of correlation study results for the 2001 and 2002 study years. This summary can serve as documentation of the generalizability of the claim of the validity of the LSAT scores for use in the law school admission process. More importantly, this study provides national longitudinal data for law schools to examine against their school-specific results to increase understanding of their own admission process. Since correlation studies are conducted for individual schools and school-specific results are reported only to the school whose data were analyzed, the results reported in this study may be used by schools as a benchmark in evaluating their own results.

The effectiveness of the LSAT score alone, undergraduate grade-point average (UGPA) alone, and the combination of these two variables for predicting first-year law school average (FYA) is evaluated. As previous studies have shown, results reported here indicate that the LSAT alone tends to be a better predictor of law school performance than is UGPA alone. The combination of LSAT and UGPA, however, continues to be superior to either predictor variable alone for predicting law school FYA. These results support the validity of the LSAT for use in the law school admission process.

Introduction

For approximately 50 years, the sponsors of the Law School Admission Test (LSAT) have offered to conduct studies of the effectiveness of the test as well as other predictors of law school performance used in the admission process. These studies are offered free of charge to participating law schools, and schools have been encouraged to avail themselves of the service. Over the years, increasingly large numbers of law schools have participated in the LSAT correlation studies. During the two-year period from 2001 to 2002, 200 different schools participated in the studies and 387 correlation studies were conducted. Among the 200 schools, 168 schools participated for both years—2001 and 2002. The summary data presented in this report are only from those 168 schools.[1] Canadian schools were excluded from this report since they did not participate in the Law School Data Assembly Service (LSDAS). Schools with non-ABA status and some schools that had grading scale changes were also excluded from this report. In addition, only fall-entering full-time students with complete data have been included.

One purpose of this study is to summarize data across schools to provide documentation of the generalizability of the claim of validity of the LSAT scores for use in the admission process. A more important purpose is to provide national longitudinal data for law schools to examine against their school-specific data to help them increase their understanding of their own admission process. Correlation studies are conducted for individual schools, and school-specific results are reported exclusively to the schools whose data were analyzed. Thus, schools know how well the test and other predictors are performing within their own admission process, but they have no benchmark against which to evaluate their results.

The correlation studies provide valuable information to LSAT score users. One task frequently assigned to those responsible for law school admission is that of identifying from large groups of law school applicants those who are most likely to succeed in law school. A limited amount of information usually is available from which to make that decision. Almost universally across all ABA-accredited law schools and English language common-law law schools in Canada, both LSAT score and undergraduate grade-point average (UGPA) are among the available data. Both are quantifiable measures that are potentially useful in making admission decisions, and many schools use this information extensively. If this (or any other) quantifiable information is relied on in the selection process, the burden is on the score user to obtain evidence that there is a relationship between the quantified variables and the outcome of interest to the admission committee—usually success in law school. The correlation studies can provide that evidence for participating schools. An additional value of the correlation studies is that they provide score users with quantifiable information about how their admission process is working and about the makeup of their entering class.

---

[1]Note that this report is largely a replication of earlier national summaries of the 1990–1992 correlation studies by Wightman (1993), the 1995–1996 correlation studies by Anthony, Harris, & Pashley (1999), the 1997–1998 correlation studies by Anthony, Duffy, & Reese (2005), the 1999–2000 correlation studies by Thornton, Suto, Anthony, & Liu (in press). The LSAT scores available at the time of the Wightman study were from the 10–48 score scale, while the three later studies analyzed LSAT scores on the 120–180 score scale.

The Criterion Variable

Academic success in law school is typically among the important outcomes that those responsible for admission would like to predict. To be sure, there exist other admission goals within individual law schools, but the LSAT purports to be useful for the limited purpose of predicting academic success, so a criterion related to academic success is the most appropriate one for validating the LSAT as a predictor. In the correlation studies, the variable used to represent academic success is first-year average (FYA) in law school. Using FYA as the criterion variable is not unique to the LSAT validity studies. A variable based on first-year grades is the most typical criterion used to validate almost all admission tests. FYA is not the only criterion that could be used, but it has several advantages that have shown it to be a useful criterion. First, it represents a composite of the academic performance of a student after a year of law school. Some of the courses taken may have been easier than others; some professors more lenient in grading than others. By using the average of all the grades received, these differences in course difficulty and grading stringency tend to average out. Second, for law students, the first-year grade-point average tends to represent basically the same curriculum for all of the students in the school. In subsequent years, different elective choices are represented in the composite average. Last, FYA data are available within a year, while other criteria might require a delay of two, three, or more years before a study can be conducted.

The Predictor Variables

Two predictor variables are used in the LSAT Correlation Studies: undergraduate grade-point average (UGPA) and LSAT score. Individual schools may use other predictors in their admission process, but these two are available for every school. The UGPA used in the correlation studies is the same as the UGPA that the Law School Admission Council provided to the law school from the Law School Data Assembly Service (LSDAS), and thus is the UGPA that was available to the law school at the time the admission decision was made. The UGPA is computed by the LSDAS, according to LSDAS procedures. Grades computed in this manner are expressed on a scale of 0.00 to 4.33.

LSAT scores available for the correlation studies reported in this study are all on the 120-180 LSAT score scale that was introduced in June 1991. The scores used in this study reflect an average of the three most recent LSAT scores. The 1995 studies were the first to be based on three years of data for LSAT scores on the 120-180 score scale. The 1995 and 1996 correlation studies, the 1997 and 1998 correlation studies, and the 1999 and 2000 correlation studies were the subjects of previous reports (Anthony, Harris, & Pashley, 1999, Anthony, Duffy, & Reese, 2005, and Thornton, Suto, Anthony, & Liu, in press). The current study is based on the 2001 and 2002 correlation study years.

The 2001 studies report data for the 1998, 1999, and 2000 fall entering classes, as available. The 2002 studies report data for the 1999, 2000, and 2001 fall entering classes. Data accumulated over three years provide more stable parameter estimates because a small number of outliers have less influence on prediction weights when the sample sizes are larger. This is particularly important for schools that use the compensatory admission model discussed later in this report. When schools allow a high test score to compensate for a low UGPA or vice versa, the range restriction in LSAT score or UGPA can change dramatically from year to year, particularly when the size of the school is relatively small. Fluctuation in range restriction causes fluctuation over time in the estimates of the raw regression weights. Aggregating data helps stabilize the weights.

Estimating Validity

The general concept of validity is a broad one, encompassing the accumulation of data to support a particular use of a test. The particular type of evidence obtained from the correlation studies is referred to as predictive validity. This is so because law school applicants' first-year law school grade-point averages are predicted from their LSAT scores and UGPAs. The statistical procedure used in these studies to predict law school performance from one or both of the prediction variables is linear regression. When a single predictor is used to predict the criterion, the prediction equation can be represented by a straight line on a graph that shows for every student a single point that represents both the score on the predictor variable (e.g., LSAT score) and the criterion score (e.g., FYA). The exact position of the line on the graph is calculated so as to minimize the (squared) distance of every point from the line. A statistic known as the correlation coefficient provides an estimate of how well the line represents the points on the graph. The correlation coefficient varies from –1 to 1. A correlation of 1 indicates a perfect positive linear relationship (high values of one variable are indicative of high values on the other variable) while a correlation of –1 indicates a perfect negative linear relationship (low values on one variable are indicative of high values on the other variable). A correlation of 0 indicates that there is no relationship between the variables being studied. When the

correlation coefficients are high, the points are close to the prediction line; when the coefficient is close to zero, there is little relationship between the points and the line. The closer the points are to the regression line (or the higher the correlation coefficient), the more accurately the predictor predicts the criterion. Figures 1, 2, and 3 show three examples of prediction lines and the relative positions of data points for some sample law school data. In Figure 1, LSAT score is the predictor, FYA is the criterion, and the correlation coefficient is .62. In Figure 2, LSAT score and FYA are again the predictor and criterion variables, and the correlation coefficient is .01. In Figure 3, LSAT score is the predictor variable, UGPA is the criterion variable, and the correlation coefficient is –0.44. The figures are provided to illustrate the relative accuracy of different sized correlation coefficients as well as to illustrate both positive and negative linear relationships.

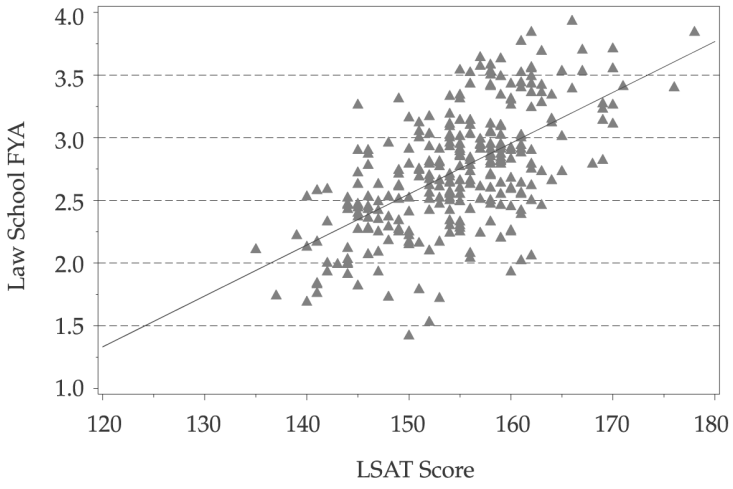

FIGURE 1. *Best regression line for LSAT/FYA data, r = 0.62*

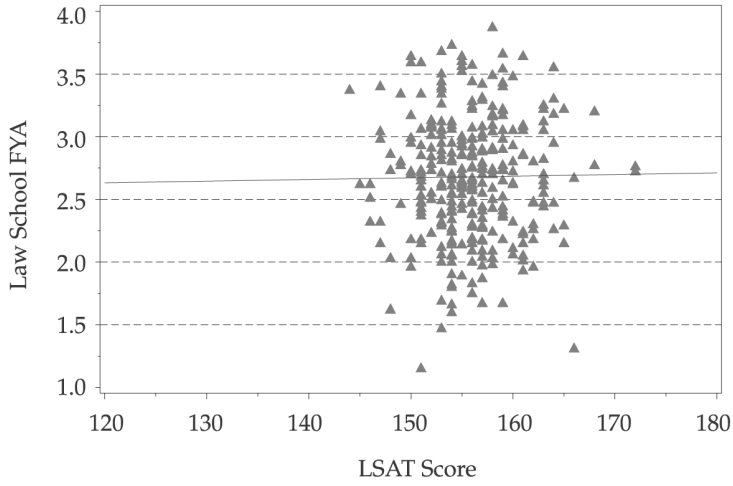

FIGURE 2. *Best regression line for LSAT/FYA data, r = 0.01*

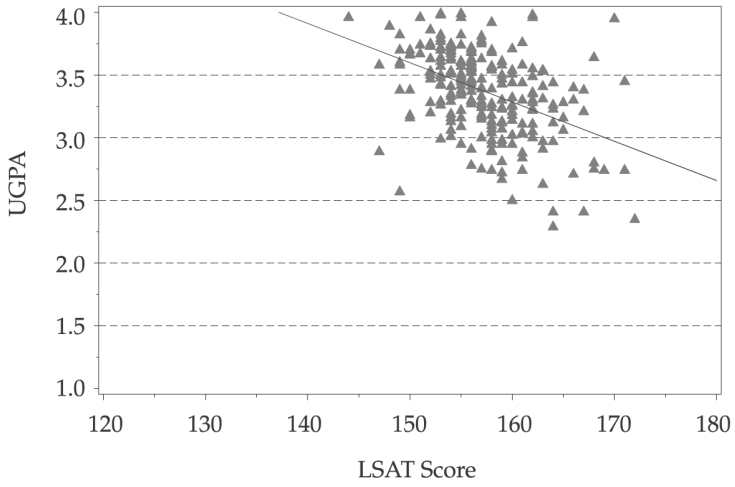

FIGURE 3. *Best regression line for LSAT/UGPA data, r = –0.44*

Test-score users should not expect to find perfect correlations between test scores and the law school performance criterion. Many factors other than the acquired academic skills measured by the LSAT contribute to academic performance. In addition, there is a certain amount of measurement error inherent in the test. When test scores and UGPAs are factors that influence the admission process, the range among admitted students becomes restricted relative to the applicant pool. The issue of restricted range draws our attention to the behavior of both the selected and unselected groups. The selected group is the students for each school that have been admitted and have an FYA. This group tends to be more homogeneous than the entire applicant pool, exhibiting less variance in both predictor variables. The unselected population of students for each school is the group that applied to a particular school during the relevant application years. Not all of these individuals have an FYA, since not all of these people were granted admittance to a particular school. A common observation of higher validity coefficients with the unselected group will also be examined in this report.

Further, because the UGPA is influenced by factors such as the leniency of the graders, the rigor of the curriculum represented by the grades, and the students' motivation and application, UGPA also should not be expected to be a perfect predictor. In fact, experience shows UGPA alone typically to be a fairly poor predictor of later academic achievement.

Results

Summary Data

Descriptive statistics for schools participating in the correlation studies for both the 2001 and 2002 study years are presented in Table 1. These data confirm that the average sample size for each correlation study is quite large. The size of the samples is primarily a consequence of including the most recent three years of student data in the study when it is available. Three years of data are not always available, and all of the studies included in this report do not include three years of accumulated data. Schools participating for the first time may only be able to provide data for one year. More commonly, the correlation study is based on a single year of data when a school introduces a grading-scale change. As evidenced by the minimum sample size of 95, even when only one year of data is available for analysis, the sample sizes are large enough to produce stable least-squares regression results. The advantage gained from using three years of data is found in the stability of the weights applied to the two predictor variables rather than in the magnitude of the correlation coefficients. The data in Table 1 show that the mean and median of both LSAT and UGPA for enrolled students showed little change from the 2001 to the 2002 correlation study years. During the same time period, the mean and median UGPAs were approximately equal. The range of average LSAT scores across schools is fairly substantial, varying from a low mean of 136.39 for the 2002 correlation study sample to a high mean of 168.66 for the same study year. The size of the range is approximately equal for both the 2001 and the 2002 correlation study samples. The range and variance of mean LSAT scores reported in Table 1 suggest that the differing characteristics of law schools are well represented in this report.

TABLE 1
Descriptive statistics for 168 schools participating in the LSAT correlation studies for 2001 and 2002

Variable	Statistic	Study Period 2001	2002
Sample Size	Mean	518.67	537.74
	Standard deviation	237.88	236.46
	Range		
	Minimum	95.00	142.00
	Maximum	1,383.00	1,410.00
	Percentiles		
	25th	356.00	374.00
	50th (median)	480.50	493.00
	75th	639.00	666.00
Mean LSAT	Mean	154.35	154.49
	Standard deviation	5.62	5.60
	Range		
	Minimum	136.78	136.39
	Maximum	168.23	168.66
	Percentiles		
	25th	150.72	150.65
	50th (median)	153.64	153.71
	75th	157.85	157.97
Mean UGPA	Mean	3.23	3.25
	Standard deviation	0.21	0.20
	Range		
	Minimum	2.69	2.72
	Maximum	3.74	3.75
	Percentiles		
	25th	3.09	3.11
	50th (median)	3.21	3.24
	75th	3.36	3.39

Correlation Coefficients

As discussed earlier, the correlation between predictor and criterion provides an indication of the usefulness of the predictor(s). The correlations presented in Table 2 show that for each of the study years, the LSAT score is a substantially better predictor of first-year performance in law school than is the UGPA. The data also show that the combination of LSAT and UGPA provides better prediction than either predictor alone. These results are consistent with findings from earlier LSAT validity summary reports (e.g., Thornton, Suto, Anthony, & Liu, in press; Anthony, Duffy, & Reese, 2005; Anthony, Harris, & Pashley, 1999; Wightman, 1993; Evans, 1982; Schrader, 1976).

TABLE 2
Summary correlations between and among predictor and criterion variables for law schools participating in 2001–2002 correlation studies: selected first-year student results

		Correlations						
				Percentiles				
Variables	Year	Mean	Standard Deviation	25	50	75	Minimum	Maximum
Zero order correlations								
LSAT/FYA	2001	0.40	0.08	0.36	0.41	0.46	0.09	0.58
	2002	0.39	0.08	0.34	0.40	0.44	0.02	0.60
UGPA/FYA	2001	0.27	0.08	0.22	0.28	0.32	0.07	0.53
	2002	0.28	0.08	0.23	0.28	0.33	0.08	0.57
LSAT/UGPA	2001	-0.02	0.13	-0.09	-0.02	0.07	-0.39	0.36
	2002	-0.03	0.13	-0.11	-0.02	0.05	-0.42	0.39
Multiple correlations								
LSAT &	2001	0.49	0.07	0.45	0.50	0.55	0.28	0.66
UGPA/FYA	2002	0.49	0.07	0.45	0.50	0.53	0.23	0.66

The size of the mean correlation coefficients is virtually identical for both years. This consistency is attributable partly to the amount of overlap in the data resulting from the accumulation of three years of data for each study by most participating schools. The size of the correlations between LSAT and FYA, and between UGPA and FYA are consistent with those reported by Thornton et al. (in press) for the 1999 and 2000 LSAT

correlations studies. That is, Thornton et al. reported median validity coefficients (correlations) of .40 for LSAT for 1999, .41 for LSAT for 2000, and .27 for UGPA for both 1999 and 2000 study years.

Stem and leaf plots provide a graphical illustration of the correlation coefficients reported in Table 2. Six separate stem and leaf plots are provided in Figures 4 and 5. Separate plots are provided for each of the predictor combinations of LSAT and UGPA combined, LSAT alone, and UGPA alone, in that order, across each row. Figures 4 and 5 show the separate distributions for the years 2001 and 2002, respectively. The stem (the left column of each plot) is the first digit of the correlation coefficient. The leaf (the row of numbers separated from the stem by a single space) shows the second digit of the correlation coefficient for each school that produced a correlation coefficient in the group defined by the stem. For example, looking at the 2001 data (Figure 4), in the stem group between .55 and .59 on the plot of correlation coefficients resulting from the combined LSAT and UGPA predictors, eight schools had correlation coefficients of .55. Thus, eight 5s are printed in the row across from the 5. Likewise, the ten 6s represent the schools that had correlation coefficients of .56, the nine 7s represent the schools that had coefficients of .57, and so on. Notice that there are two rows that begin with a 5. The leaves in one row range from .50 to .54; the leaves in the other row range from .55 to .59. This pattern holds for every pair of rows in the figures. The right-hand column of each plot shows the number of schools represented in each stem and leaf (row) of the plot. If the stem and leaf plot is rotated 90 degrees, it can be viewed as a histogram of the distribution of correlation coefficients. Examination of the stem and leaf plots reveals that the coefficients are not widely variable across schools and the histograms are very peaked. Looking across the three plots within a single year affords an informative picture of the usefulness of the different predictors in providing information about first-year performance in law school. The correlation coefficients derived from the combined predictors are concentrated at the highest section of the plot because the combined predictors produce the highest correlation with FYA. Moving from top to bottom on the page, the concentration of correlation coefficients steps down the stem axis, indicating slightly lower correlations resulting from using LSAT alone, and considerably lower correlations using UGPA alone. The step-down pattern is consistent across each of the two years of studies, and, of course, is consistent with the summary data reported in Table 2.

Individuals might find it of interest to locate their own school's correlation coefficient on each of the three stem and leaf plots for the relevant study year. By marking or circling their own school's coefficient, the individual will have an indication of how well the predictor or combined predictors are working for their school relative to other law schools. This information needs to be evaluated in context with other pertinent information, particularly the variability of the predictor(s) in the entering class(es) upon which the analyses are based and the correlation between LSAT and UGPA for members of the analysis class. The impact of each of these factors is discussed in later sections of this report.

```
Variable = LSAT and UGPA combined
  Stem Leaf                                                                  #
6 56                                                                         2
6 001111122                                                                  9
5 555555556666666666777777778888889                                         34
5 000000000011111222222222333333333344444                                   40
4 555555666677777777777778888888888899999999                                43
4 000000011111111222333333344                                               28
3 555578999                                                                  9
3 2                                                                          1
2 88                                                                         2
2
1
1
0
0

Variable = LSAT alone
  Stem Leaf                                                                  #
6
6
5 558                                                                        3
5 00011222233444                                                            14
4 555555555666666677777778888999999999                                      35
4 000000000111111111222222233333333333344444                               42
3 555566666666677777788888899999999999999                                   37
3 011111222333444444                                                        18
2 677788888999                                                              12
2 33334                                                                      5
1
1 4                                                                          1
0 9                                                                          1
0

Variable = UGPA alone
  Stem Leaf                                                                  #
6
6
5
5 3                                                                          1
4 59                                                                         2
4 00111112224                                                               11
3 55555567788899                                                            14
3 000000000001111111111222222222223333444444                               44
2 555555555566677777777777888888888899999999                               42
2 000001111112222233334444444                                              26
1 56666777778889999                                                         17
1 0012344                                                                    7
0 7789                                                                       4
0
```

FIGURE 4. *Stem and leaf plots for prediction of FYA from LSAT score and UGPA in combination and alone, 2001 data*

Note. Multiply Stem.Leaf by 0.1

```
Variable = LSAT and UGPA combined
  Stem Leaf                                                                    #
6 6                                                                            1
6 0233                                                                         4
5 55555566666667788889999999                                                  27
5 000000000000011111111111112222222222333333333444444444                     56
4 55555556666666677777777777777788888999999999                                43
4 0000001111222222233444                                                      22
3 5557889999                                                                  10
3 244                                                                          3
2 5                                                                            1
2 3                                                                            1
1
1
0
0

Variable = LSAT alone
  Stem Leaf                                                                    #
6
6 0                                                                            1
5 7                                                                            1
5 000112223344                                                                12
4 555566666677777788888899                                                    24
4 000000000011111111111122222223333333344444444                              47
3 55555555566666677777778888888899999999999                                  40
3 000001111233333334444444                                                    24
2 666778899999                                                                12
2 0224                                                                         4
1 6                                                                            1
1
0
0 24                                                                           2

Variable=UGPA alone
  Stem Leaf                                                                    #
6
6
5 7                                                                            1
5
4
4 00012233334                                                                 11
3 5556667777888999                                                            16
3 000000111111111111222222233333333344444444                                 43
2 55555555555566666666777777888888889999999999999                            46
2 000011112222222222333333344444444                                           32
1 666778888                                                                    9
1 122244                                                                       6
0 8999                                                                         4
0
```

FIGURE 5. *Stem and leaf plots for prediction of FYA from LSAT score and UGPA in combination and alone, 2002 data*
Note. Multiply Stem.Leaf by 0.1

Factors Influencing the Magnitude of the Correlation Coefficients

The correlation coefficients reported in Table 2 most likely are an underestimate of the true validity of the test. That is, the correlations are based on LSAT scores, UGPAs, and FYAs only for those students who were accepted to and attended the studied law school. Most applicants with low test scores and low UGPAs are not admitted and thus there are no first-year grades for them. As a consequence, those applicants cannot be included in the study. Because there is less variability in the scores of admitted students than in the scores of all applicants, correlations are smaller than they would have been had the class been admitted randomly from the total applicant pool. Thus, correlations obtained from matriculated students tend to underestimate the true validity of the test. Even so, they are the best information that we have available, and even as underestimates they are quite reputable.

Table 3 presents estimated and actual correlation coefficients for the unselected population of applicants. As expected, the validity coefficients are consistently higher for the unselected group than they are for the selected group represented in Table 2. For example, when comparing the selected and unselected groups in the 2001 and 2002 study years, the validity coefficients for LSAT/FYA increased .13 and .14, respectively. For both study years, the minimum and maximum values of the correlation coefficients for each combination of variables tend to be stable. For example, in 2001 the range of the correlation coefficient for UGPA/FYA was .07 to .54. In 2002, the range was .09 to .57. For each combination of variables, the mean and median values of the validity coefficients tend to be similar, if not equal. In both 2001 and 2002, the mean correlation coefficient of .27 was observed for both mean and median for LSAT/UGPA.

TABLE 3

Summary correlations between and among predictor and criterion variables for law schools participating in 2001–2002 correlation studies: unselected results

		Correlations						
			Standard	Percentiles				
Variables	Year	Mean	Deviation	25	50	75	Minimum	Maximum
LSAT/FYA*	2001	0.53	0.07	0.48	0.55	0.58	0.31	0.69
	2002	0.53	0.07	0.48	0.53	0.57	0.30	0.69
UGPA/FYA*	2001	0.31	0.09	0.26	0.32	0.37	0.07	0.54
	2002	0.32	0.08	0.27	0.32	0.37	0.09	0.57
LSAT/UGPA**	2001	0.27	0.06	0.23	0.27	0.31	0.07	0.44
	2002	0.27	0.05	0.23	0.27	0.30	0.11	0.44

*Adjusted for restriction of range (see equation in Appendix A)

**Based on all applicants

In addition to the problem of reduced variability, matriculated students include some who are admitted as a result of special consideration. That is, some students with low test scores or low UGPAs are admitted to law school, but usually they are not typical of the low-scoring applicants who are rejected. Instead, they are admitted because the school has some other evidence of their ability to do well in law school. This practice frequently is referred to as a compensatory admission model. For example, a compensatory model allows a high LSAT score to compensate for a low UGPA and conversely, a high UGPA to compensate for a low LSAT score when making admission decisions. One way to determine whether a compensatory model for LSAT scores and UGPAs is in place in a particular law school is to look at the correlation between LSAT score and UGPA. In a random group of applicants, this correlation would be fairly high, indicating that applicants with high LSAT scores also had high UGPAs, while applicants with low LSAT scores also had low UGPAs. When a compensatory model is used, the correlation between LSAT and UGPA frequently is negative because a large number of students with high LSAT scores have low UGPAs and vice versa.

The average correlations between the LSAT and UGPA reported in Table 2 are close to zero. Across the two years, the correlations range from –.42 to .39, suggesting that a substantial number of law schools are to some degree employing a compensatory admission model. The data presented in Table 4 show the effect of employing a compensatory model on the estimates of the validity coefficients. Schools that rely heavily on a compensatory model tend to have negative correlations between LSAT and UGPA and are represented in the rows labeled "Less than 0" in Table 4. Over half of the schools fell into this category for each of the years reported. The data show that the correlations for either predictor alone and for the two predictors combined are consistently higher for schools where the correlation between LSAT and UGPA is positive. For example, looking at the 2002 correlation study data, the average correlation between LSAT and FYA is .36 for schools that have a negative correlation between LSAT and UGPA, but it increases to .47 for schools that show a positive correlation greater than .2 between those variables. A similar pattern exists for UGPA alone and for LSAT and UGPA combined. The pattern is evidenced across both years.

TABLE 4
Average correlations of LSAT scores and UGPA with FYA in law schools grouped by the correlation between LSAT score and UGPA

		Predictor Variables		
Year	Correlation of LSAT with UGPA	LSAT & UGPA Combined	LSAT Alone	UGPA Alone
2001	Less than 0	0.47	0.37	0.23
	(number of schools)	(93)	(93)	(93)
	0.0 to 0.2	0.52	0.43	0.32
	(number of schools)	(68)	(68)	(68)
	Greater than 0.2	0.56	0.49	0.38
	(number of schools)	(7)	(7)	(7)
2002	Less than 0	0.46	0.36	0.25
	(number of schools)	(96)	(96)	(96)
	0.0 to 0.2	0.52	0.43	0.31
	(number of schools)	(64)	(64)	(64)
	Greater than 0.2	0.56	0.47	0.41
	(number of schools)	(8)	(8)	(8)

One observation of note from the data presented in Table 2 is that the range of correlation coefficients for any of the prediction models varies substantially from law school to law school. For example, the correlations between LSAT and FYA vary from a low of .02 to a high of .60. The cause for this amount of variation in what might be expected to be a stable environment was studied extensively by Linn (e.g., Linn, 1982; Linn, Harnish, & Dunbar, 1981). In analyzing the validity summary data reported by Schrader (1976), Linn determined that as much as 34% of the variance in observed validities could be predicted from observed standard deviations and variances of the LSAT. That is, the larger the variation (range) in the predictor, the higher the correlation with the criterion. Because so much variation is observed in the correlation coefficients for the validity studies summarized in this report, Linn's procedures were replicated using the 2001–2002 correlation study data. For each study year, the multiple correlations between the observed validities and observed standard deviations and variances of the LSAT were calculated. (Recall that the variance is equal to the square of the standard deviation.) Similarly, the same calculations were made using observed standard deviations and variances of UGPA as well as the combined LSAT and UGPA standard deviations and variances. The results are presented in Table 5. The amount of variance in observed validities predicted by the LSAT is slightly lower than was reported by Linn. For the 2001 data, 30% of the variance (.55) is accounted for, and for the 2002 data, 27% of the variance (.52) is accounted for. For both the LSAT alone and for the LSAT and UGPA combined, there is a substantial relationship between variability among validity coefficients across law schools and variability in the predictors within law schools. In contrast, the variance in UGPA accounts for approximately 14% of the variance in the UGPA validity coefficients. Even so, the relationship between validity and UGPA variability is larger than the .05 correlation reported by Linn and Hastings (1983).

TABLE 5
Multiple correlations of validity coefficients with standard deviations and variances of predictor variables

	Predictor		
Year	LSAT Alone	UGPA Alone	LSAT & UGPA Combined
2001	0.55	0.37	0.55
2002	0.52	0.38	0.54

Cross Validation Studies

A primary purpose for conducting validity studies for most schools is to obtain the best possible prediction weights so that they can be applied to the application credentials of the next year's applicant pool to aid in the decision process. That is, data from past experiences are used to make future predictions. When the results from the correlation studies are being used in this way, the most relevant question to ask is how well do the equations from previous first-year classes predict the performance of future first-year classes. To answer this question, several cross validation studies were conducted. Specifically, the prediction equations

calculated from the 2001 correlation studies (i.e., the 1998, 1999, and 2000 first-year class data) were used to predict an FYA for each member of the 2001 fall entering first-year class. Then, the correlation between the predicted FYA and the actual FYA earned by the members of the 2001 fall entering class was calculated. These calculations were performed separately for each school, using each school's unique least-squares prediction model. The results from these cross validations are presented in Table 6. The results from these analyses are nearly identical to the correlation coefficients reported in Table 2. When the equations from the immediately preceding year are used to predict FYA, the correlation between predicted and actual FYA exceeds .4 for more than 82% of the schools. These results are consistent with those reported by Evans (1982) for the 1977–79 studies, when 73%, 74%, and 82%, respectively, of the schools exceeded the correlation coefficient value of .4. The results shown by Wightman (1993) for the 1990–92 studies reported 86% of the schools exceeding the correlation coefficient of .4. The 2001–2002 results are slightly higher than those reported by Thornton et al. (in press) for the 1999-2000 studies when more than 81% of the schools exceeded the correlation coefficient of .4.

TABLE 6
Cross validated multiple correlations for 2001 prediction equations using 2002 class data

Correlations between Actual FYA and Predicted FYA	2001 Equations applied to Fall 2001 Class
Mean	0.48
Standard deviation	0.09
Range	
Minimum	0.15
Maximum	0.66
Percentiles	
25th	0.43
50th	0.49
75th	0.54
Percentage GT .4	82.74

Accounting for Variance

The data in Table 2 show that the best model for predicting FYA is consistently the model that combines LSAT and UGPA, where best model is defined as the model that provides the highest correlation between the predictors and the criterion. On average, the LSAT accounts for more of the variance predicted by the model than does UGPA, although there are some schools for which this is not the case. As shown in Table 7, the mean variance accounted for by LSAT is approximately 58% and, thus, the mean variance accounted for by UGPA is approximately 42%. Another way of interpreting the percentage of variance accounted for is in terms of the relative weights of the two predictors. That is, the LSAT is weighted 58% and the UGPA is weighted 42% on average in order to obtain optimal prediction of FYA. The percent of variance accounted for by LSAT and by UGPA can vary considerably for individual schools. The variability of LSAT scores and UGPA, the correlation between UGPA and LSAT, and the amount of variability in the FYAs all influence the amount of variance that will be accounted for by the two predictor variables in the model that provides optimal prediction of FYA.

TABLE 7
Summary of percentage of variance accounted for by predictor variables in multiple regression prediction equations

				Range		Percentiles		
Variable	Year	Mean	Standard Deviation	Minimum	Maximum	25	50	75
LSAT	2001	58.89	7.11	30.70	72.96	54.55	59.69	64.14
	2002	57.64	7.68	20.87	73.86	52.99	57.90	62.50
UGPA	2001	41.11	7.11	27.04	69.30	35.86	40.31	45.46
	2002	42.36	7.68	26.14	79.13	37.50	42.10	47.02

Trends Over Time

The data presented in Tables 2 through 7 suggest that the results from the validity studies conducted for each of the 168 law schools that participated in the two most recent study cycles are consistent from year to year. Analyses to more specifically examine data consistency across years are presented in Table 8. The average difference between validity coefficients of studies conducted for 2001 and 2002 is very close to zero. Additionally, the distribution of differences is fairly tight, as evidenced by the difference values that mark the 25th and 75th percentiles. Again, these results are somewhat influenced by the inclusion of three years of

data. For each subsequent year, the oldest data year is dropped and the most recent is added. Although this data design will minimize the influence of individual anomalies, systematic shifts in the data over time would be evidenced should they occur. The data reported in this study do not indicate changes in the validity of either predictor alone or in the combined predictors for the national data reported for the years 2001 and 2002.

TABLE 8
Average size of year differences in validity coefficients by type of prediction model

Years	Differences	LSAT & UGPA Combined	LSAT Alone	UGPA Alone
2001–2002	Mean	0.01	0.01	0.00
	Standard deviation	0.03	0.04	0.04
	Percentiles			
	25th	–0.01	–0.01	–0.03
	50th	0.01	0.01	–0.01
	75th	0.03	0.04	0.02

Conclusion

This two-year national summary of the LSAT correlation studies conducted in the years 2001 and 2002 lends continued support for the validity of the LSAT scores in the law school admission process. Major findings from this study are summarized as follows:

- The combination of LSAT and UGPA are useful predictors of academic performance in the first year of law school. The average multiple correlation between first-year grades in law school and the combined predictors of LSAT and UGPA is .49 for both 2001 and 2002. This value is high and consistent with that reported previously. As has always been the case, these combined predictors continue to be superior to either predictor alone for predicting FYA.

- LSAT alone continues to be a better predictor of law school performance than is UGPA alone. The median validity for LSAT alone is .41 (2001) and .40 (2002), compared with .28 for UGPA alone for both years.

- When schools are grouped by the correlation between LSAT and UGPA, the validity coefficients increase when the correlations between the predictors increase. This relationship provides some indication of the impact of the restriction of range resulting from using only matriculated students on the estimates of validity, particularly in the presence of a compensatory admissions model.

- A substantial amount of the variability in validity coefficients obtained among different law schools is directly attributable to the amount of variation in LSAT scores and UGPAs in the data used to estimate the validity.

- Cross-validation studies support the use of regression equations based on previous first-year classes to predict future performance of law school applicants.

An additional important outcome of the current study is that the data show no deterioration in the validity of the LSAT as a predictor of first-year performance in law school since the conduct of the validity summary study by Evans (1982) that looked at results from correlation studies carried out in 1977 through 1979. During this time, the UGPA also has shown no change in its predictive power. Thus, the relative importance of the two remains essentially unchanged.

References

Anthony, L. C., Harris, V. F., & Pashley, P. J. (1999). *Predictive validity of the LSAT: A national summary of the 1995–1996 correlation studies* (LSAT Technical Report No. 97-01). Newtown, PA: Law School Admission Council.

Anthony, L. C., Duffy, J. R., & Reese, L. M. (2005). *Predictive validity of the LSAT: A national summary of the 1997–1998 correlation studies* (LSAT Technical Report No. 99-01). Newtown, PA: Law School Admission Council.

Evans, F. R. (1982). Recent trends in law school validity studies. In Law School Admission Council, *Reports of LSAC sponsored research: Volume IV, 1978–1983* (LSAC Report No. LSAC-82-1, pp. 347–361). Newtown, PA: Law School Admission Council.

Linn, R. L. (1982). Admissions testing on trial. *American Psychologist, 37*, 279–291.

Linn, R. L., Harnish, D. L., & Dunbar, S. B. (1981). Validity generalization and situational specificity: An analysis of the prediction of first-year grades in law school. *Applied Psychological Measurement, 5*, 281–289.

Linn, R. L., & Hastings, C. N. (1983). *A meta analysis of the validity of predictors of performance in law school*. In Law School Admission Council, *Reports of LSAC sponsored research: Volume IV: 1978-1983* (LSAC Report No. LSAC-83-1). pp. 507–544). Newtown, PA: Law School Admission Council.

Schrader, W. B. (1976). Summary of law school validity studies, 1948–1975. In Law School Admission Council, *Reports of LSAC Sponsored Research: Volume III, 1975–1977* (LSAC Report No. LSAC-76-8, pp. 519–550). Princeton, NJ: Law School Admission Council.

Thornton, A. E., Suto, D. A., Anthony, L. C., & Liu, M. (in press). *Predictive validity of the LSAT: A national summary of the 1999-2000 correlation studies* (LSAT Technical Report No. 01-02). Newtown, PA: Law School Admission Council.

Wightman, L. F. (1993). *Predictive validity of the LSAT: A national summary of the 1990-1992 correlation studies* (Research Report No. 93-05). Newtown, PA: Law School Admission Council.

Appendix A

Only the admitted applicants who enroll full-time and complete their first year will have FYA data. To compute the unselected correlation coefficients with FYA, a statistical adjustment must be applied to the available data on the selected group of students who matriculate. We use the following:

$$r_{XY}^{*} = \frac{1}{\sqrt{1 + \frac{\sigma_{X^*}^2}{\sigma_X^2}\left(\frac{1}{r_{X^*Y^*}^2} - 1\right)}}$$

where r_{XY}^{*} is the estimated correlation coefficient for the X and Y after correcting for the selection effect, $\sigma_{X^*}^2$ is the variance of variable X in the selected sample, σ_X^2 is the variance of X in the unselected sample, and $r_{X^*Y^*}$ is the correlation between X and Y in the selected sample.