LSAT TECHNICAL REPORT SERIES

Analysis of Differential Prediction of Law School Performance by Gender Based on 2011–2014 Entering Law School Classes

Tammy J. Trierweiler
Nazia Rahman

Law School Admission Council
LSAT Technical Report 17-02
March 2017

A Publication of the Law School Admission Council

The Law School Admission Council (LSAC) is a nonprofit corporation that provides unique, state-of-the-art products and services to ease the admission process for law schools and their applicants worldwide. Currently, 222 law schools in the United States, Canada, and Australia are members of the Council and benefit from LSAC's services. All law schools approved by the American Bar Association are LSAC members. Canadian law schools recognized by a provincial or territorial law society or government agency are also members. Accredited law schools outside of the United States and Canada are eligible for membership at the discretion of the LSAC Board of Trustees; Melbourne Law School, the University of Melbourne is the first LSAC-member law school outside of North America. Many nonmember schools also take advantage of LSAC's services. For all users, LSAC strives to provide the highest quality of products, services, and customer service.

Founded in 1947, the Council is best known for administering the Law School Admission Test (LSAT®), with about 100,000 tests administered annually at testing centers worldwide. LSAC also processes academic credentials for an average of 60,000 law school applicants annually, provides essential software and information for admission offices and applicants, conducts educational conferences for law school professionals and prelaw advisors, sponsors and publishes research, funds diversity and other outreach grant programs, and publishes LSAT preparation books and law school guides, among many other services. LSAC electronic applications account for nearly all applications to ABA-approved law schools.

This study is published and distributed by LSAC.

Table of Contents

Executive Summary

The Law School Admission Council (LSAC) has carried out annual predictive validity studies, also called LSAT Correlation Studies, since the Law School Admission Test (LSAT) was first administered. These studies are geared toward evaluating and ensuring the effectiveness and validity of LSAT scores for use in the law school admission process. In conjunction with these predictive validity studies, LSAC also conducts differential validity and differential prediction studies on the LSAT to ensure that the test is fair across gender subgroups. The purpose of this report is to summarize the results of the 2012–2015 LSAT Correlation Studies, which are based on the 2011–2014 entering law school classes of participating schools, in a differential validity framework. The results presented serve to document and support the validity of LSAT scores for use in the law school admission process.

This study examined results for male and female law students. Data were analyzed from 158 law schools. Validity coefficients and prediction equations using LSAT score with first-year average (FYA), undergraduate grade point average (UGPA) with FYA, and the combination of both LSAT score and UGPA with FYA were calculated and evaluated for each gender subgroup.

Results of analyses indicate that the validity coefficients calculated for each gender subgroup were very similar to one other. In addition, the amount of over- or underprediction for each predictor variable alone and for the combination of both predictor variables in predicting law school FYA was quite small. Still, the combination of both LSAT score and UGPA as predictors provided the least amount of over- or underprediction for both male and female law students compared to the use of either predictor alone. Overall, results do not suggest that the use of LSAT score alone or the combination of LSAT score and UGPA contributes to unfair admission decisions for either male or female law school applicants.

At least two caveats should be remembered when evaluating the results of this study. First, only differences in average predicted performance were analyzed. That is, the performance of individuals within a subgroup whose FYAs are overpredicted on average may still be underpredicted, and vice versa. Second, differential prediction is only one aspect of an overall construct validity evaluation. Other aspects of validity should also be considered when deciding whether the use of any test scores is valid.

Introduction

The Law School Admission Council (LSAC) has carried out annual predictive validity studies, also called LSAT Correlation Studies, on the Law School Admission Test (LSAT) since the test was first administered. The primary purpose of these correlation studies is to evaluate and ensure the effectiveness and predictive validity of the LSAT for use in the law school admission process. Predictive validity refers to the degree to which a variable predicts measurement of a construct at a later time. For example, the relationship between scores on the LSAT administered to prospective law students and their later first-year law school performance may be analyzed to evaluate the predictive validity of the LSAT.

In addition to providing evidence of the overall predictive validity of the LSAT, it is important to ensure that the inferences drawn from test scores are supported with validity evidence regarding fairness, as stakes associated with admission decisions are high. Although fairness is a social construct, in the law school admission process, fairness is related to judgments about how the test scores are interpreted and used in the process of evaluating and selecting students.

Two terms often used in studies evaluating test fairness in the admission process are differential validity and differential prediction. *Differential validity* exists if the magnitude of the validity coefficient differs for one particular subgroup versus another (Linn, 1978). For example, if the magnitude of the correlation between LSAT scores and law school first-year averages (FYAs) varies across relevant subgroups (i.e., male versus female law students) this could suggest the presence of differential validity by gender subgroup.

Differential prediction focuses on subgroup differences in the prediction of an outcome. More specifically, differential prediction investigates whether the prediction model developed based on all students has the same meaning across different subgroups. For example, if it is assumed that a measure (e.g., LSAT score, undergraduate grade point average [UGPA]) predicts performance in law school, then if the prediction of performance is substantially different for one subgroup versus another, this suggests that prediction may have different meanings for each of the subgroups. If one subgroup of the applicant population experiences either significantly more overprediction (i.e., average predicted FYA greater than average observed FYA) or significantly more underprediction (i.e., average predicted FYA less than average observed FYA) than some other subgroup, then differential prediction is said to occur. The need to investigate differential prediction for relevant subgroups is directly

addressed by the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014, p. 66).

Indeed, these questions are not new to research sponsored by LSAC, nor are they unique to the LSAT or to the law school admission process. Several studies using LSAT data to investigate questions of differential validity and differential prediction across gender subgroups have been conducted by LSAC (Norton, Suto, & Reese, 2013; Suto, Norton, & Reese, 2007, 2010). Differential prediction has also been the subject of research studies for other admission testing programs such as the SAT (e.g., Mattern, Patterson, & Kobrin, 2012; Shaw, Kobrin, Patterson, & Mattern, 2012), the ACT (e.g., Noble, 2003), and the GRE (e.g., Burton & Wang, 2005).

The purpose of the current study is to address questions of differential validity and differential prediction for gender subgroups, using data from the 2011–2014 first-year law school classes. The present study is part of an ongoing monitoring effort designed to address the following question: Do any of the traditional predictors of first-year law school performance—LSAT score, UGPA, or the combination of both—result in differential prediction on the basis of gender?

Methods

Sample

The sample used in this study is drawn from LSAT Correlation Study data for the 2011–2014 entering law school classes of ABA-approved schools. Of these schools, 158 participated in the 2012–2015 LSAT Correlation Studies. Gender categories were based on self-report, and all 158 schools had a sufficient number of male and female students for data to be retained for the purposes of the current study.[1] Data from each of the 158 participating ABA-approved law schools were combined across the study years to ensure stability in the analysis and to increase the representation of the law school. The total sample included in the analyses in this study was $N = 99{,}025$.

[1]Canadian law schools were excluded from this report because they did not participate in LSAC's Credential Assembly Service (CAS).

Variables

The following variables were included in this study:

First-year average (FYA). The FYA is the average grade earned by each full-time fall-entering student in their first year of law school. As different law schools use different scales for first-year grades, FYA values were transformed to a scale with a mean of 50 and a standard deviation of 10. Student FYA data were aggregated within schools, and results are reported on this transformed scale.

Undergraduate grade point average (UGPA). The average grade earned by each student during his or her undergraduate study is computed by LSAC's Credential Assembly Service (CAS) and is expressed on a scale from 0.00 to 4.33. Student UGPAs were aggregated within schools.

LSAT score. LSAT score data were obtained from the 2012–2015 LSAT Correlation studies. Total LSAT scores are reported on a 120–180 scale, and LSAT scores were aggregated within school across test takers. If a test taker took the LSAT more than once, the average of all reportable LSAT scores for that test taker was used.

Self-reported gender. Gender designation categories for students were based on self-report, and only those students who selected a gender designation were included.

Analysis Methods

As the primary aim was to evaluate the fairness and appropriateness of using LSAT score and UGPA, either separately or in combination, as predictors of law school performance, two different types of analyses were carried out. First, validity coefficients (i.e., correlations) between each predictor and the criterion (FYA) were calculated and evaluated for each gender subgroup. As mentioned above, if relationships between predictors and FYA differ substantially between subgroups, this could be an indication of differential validity. Second, regression models were developed using combined data across all of the student data (male and female), and the prediction results were evaluated to assess whether the regression equations resulted in differential prediction for either gender subgroup.

Differential Validity

In this study, separate bivariate correlations between LSAT scores and FYA, as well as between UGPA and FYA, were evaluated and compared for male and female law students. A correlation describes the linear relationship between two variables. Correlation values can range in value from –1 to 1: A positive correlation indicates that high values on one variable are indicative of high values on the other variable, a negative correlation indicates that low values on one variable are indicative of high values on the other variable, and a correlation of 0 indicates that there is no relationship between the two variables under study.

Bivariate correlations were first calculated at the school level and then averaged within gender subgroups for comparison purposes. In addition, multiple regression analysis was used to obtain a multiple correlation value that indicates the correlation between the predicted FYA and the observed FYA when LSAT score and UGPA were used as combined predictors. As with the bivariate analyses, multiple regression analyses were first conducted at the school level, and the resulting multiple correlations were averaged within gender subgroups for comparison.

Differential Prediction

As mentioned above, differential prediction investigates whether a common prediction model is fair to all subgroups of students. In this study, three separate least-squares regression equations were used to predict FYA. These regression equations included using LSAT score alone as a predictor of FYA, UGPA alone as a predictor of FYA, and using LSAT score and UGPA as combined predictors of FYA. Differences between the predicted FYA and the observed FYA were evaluated for each model by subgroup. Because each school has a distinct grading scale, before running the regression analysis, a conversion was made to allow comparisons across law schools and to preserve the confidentiality of the school-level data. As described above, FYAs were converted to a scale where the mean for the total group of students at each school was set to 50 and the standard deviation to 10.

For each regression equation, in order to calculate mean differences, average observed FYAs for each subgroup were subtracted from the subgroup's predicted FYAs within each school. A resulting negative difference indicates that the respective regression equation underpredicts the average performance of a subgroup in a law school, while a positive value indicates that the regression equation overpredicts the mean performance of a subgroup in a law school.

Results

The results from this study are presented in three parts. The first part includes descriptive data for each gender subgroup. The second part reports the validity coefficients between the predictor variables and FYA for each gender subgroup. In the third part, the results of applying the prediction equations derived using the total group data are reported for the gender subgroups.

Descriptive Statistics

Descriptive statistics for the sample of students within the law schools used in this study are presented in Tables 1–5 and Figure 1. These data provide information about the number and proportion of male and female students and the size of the gender subgroups among the law schools included in this study. The tables and figure also allow for the comparison of LSAT score, UGPA, and FYA for male and female first-year law students.

Table 1 describes the overall gender subgroup breakdown for schools that participated in the 2012–2015 LSAT Correlation Studies. Of the 102,643 students at the 166 schools represented across the 4 years, 53,769 (52.4%) were male and 48,874 (47.6%) were female.

TABLE 1
Number and percentage of male and female first-year law school students among schools that were included in the 2015 LSAT Correlation Studies

			Male Students		Female Students	
Entering Class	Total	No. of Schools	*N*	%	*N*	%
2011	27,281	144	14,588	53.5%	12,693	46.5%
2012	25,071	147	13,187	52.6%	11,884	47.4%
2013	25,121	158	13,077	52.1%	12,044	47.9%
2014	25,170	166	12,917	51.3%	12,253	48.7%
Pooled data	102,643	166	53,769	52.4%	48,874	47.6%

Table 2 describes the gender breakdown for those schools that participated in the 2012–2015 LSAT Correlation Studies and met the criteria for inclusion in this study. The sample for the analyses used in this study included 158 schools and a total of 99,025 students. The proportion of male and female students was very similar across years and is also in line with demographic percentages reported by the Law School Admission Council and the American Bar Association (ABA, 2016).

TABLE 2
Number and percentage of male and female first-year law school students among schools that were included in this study

			Male Students		Female Students	
Entering Class	Total	No. of Schools	*N*	%	*N*	%
2011	26,467	137	14,178	53.6%	12,289	46.4%
2012	24,360	141	12,810	52.6%	11,550	47.4%
2013	24,103	150	12,543	52.0%	11,560	48.0%
2014	24,095	158	12,370	51.3%	11,725	48.7%
Pooled data	99,025	158	51,901	52.4%	47,124	47.6%

Table 3 provides a distribution of law schools by percentage of gender subgroup enrollment across the 158 law schools, revealing a slightly greater percentage of law schools with a higher concentration of male students than female students.

TABLE 3
Distribution of law schools by percentage of gender subgroup enrollment

	Male Students		Female Students	
% Subgroup Enrollment	No. of Schools	%	No. of Schools	%
≤ 30	0	0.0%	0	0.0%
31–40	3	1.9%	19	12.0%
41–50	50	31.6%	95	60.1%
51–60	95	60.1%	41	25.9%
61–70	10	6.3%	3	1.9%
≥ 71	0	0.0%	0	0.0%

Table 4, summarizing size categories by gender subgroup for the included schools, reveals that the representation of law schools having an average of 400–499 male students is nearly twice the number of law schools with the same representation of female law students; conversely, the female subgroup significantly outnumbers the male subgroup in the 100–199 size category. For the other size categories, however, the representation is very similar across gender subgroups.

TABLE 4
Summary of the number of law schools included in this study, by average size of gender subgroup

	Size of Gender Subgroup					
Gender Subgroup	<100	100–199	200–299	300–399	400–499	≥ 500
Male	8	21	49	35	31	14
Female	7	37	49	33	17	15

Table 5 provides descriptive statistics by gender subgroup for the variables used in this study. Results indicate that the mean LSAT score and mean FYA are both slightly higher for the male subgroup, while the mean UGPA is slightly higher for the female subgroup.

TABLE 5
Descriptive statistics of study variables for law schools included in this study

Gender Subgroup	No. of Schools	*N*	Mean LSAT Score		Mean UGPA		Mean FYA	
			M	*SD*	*M*	*SD*	*M*	*SD*
Male	158	51,901	155.99	5.18	3.30	0.19	50.11	0.58
Female	158	47,124	154.70	5.07	3.43	0.17	49.86	0.66

Figure 1 presents differences between mean LSAT scores, UGPAs, and FYAs for male and female students. Mean differences were calculated for each subgroup separately within each school. The *x*-axis reflects the mean score difference between male and female students, and the *y*-axis indicates the number of schools displaying each observed mean difference. The vertical dashed line indicates a mean difference value of zero for reference. Note that across the graphs for different variables (e.g., LSAT score, UGPA, and FYA) within the figure, the scale represented in the horizontal axes will change to reflect scale differences. Figure 1 illustrates the results presented in Table 5. As the figure shows, male students outperformed female students in relation to both mean LSAT scores and mean FYA in law school, whereas female students outperformed male students in relation to mean UGPA.

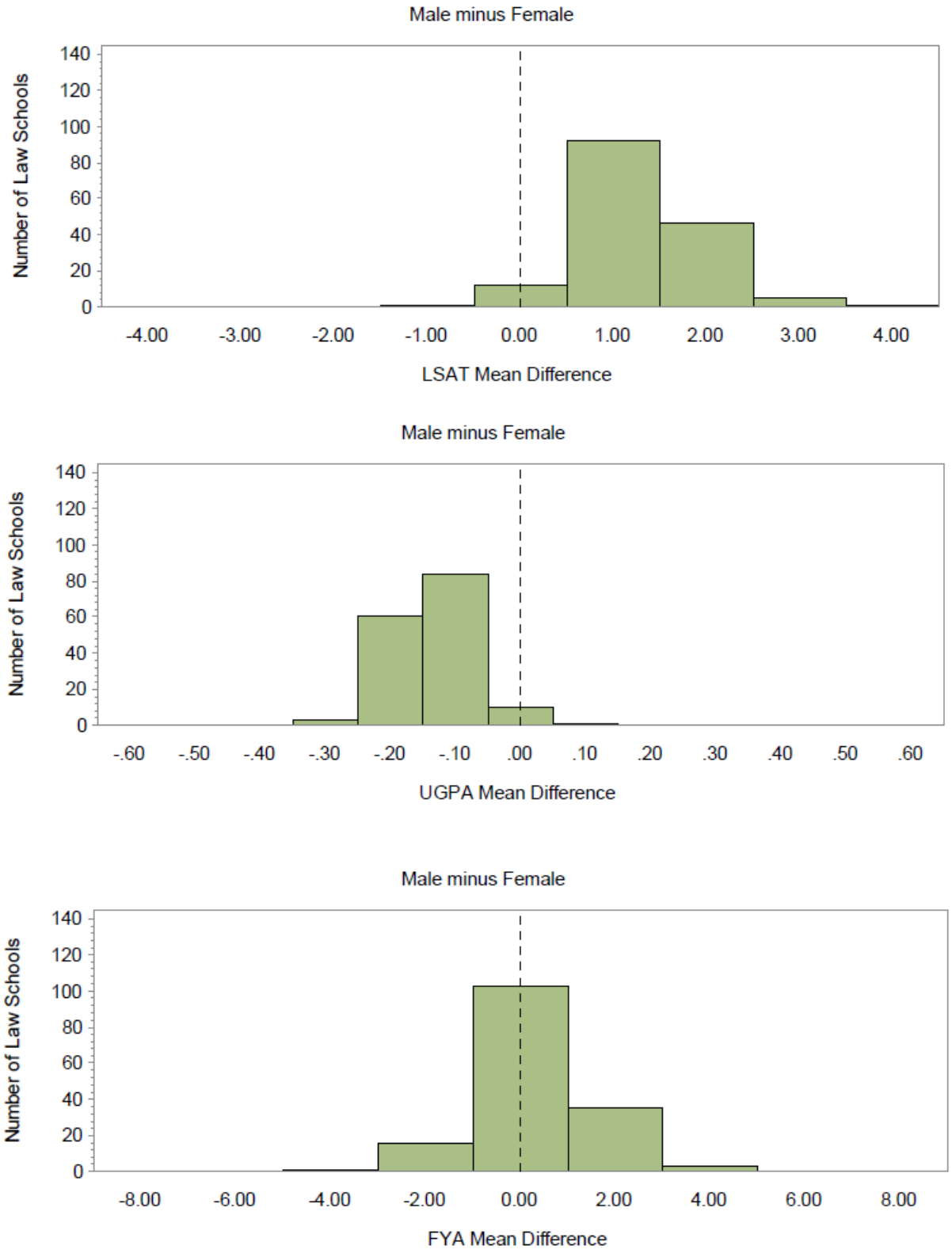

FIGURE 1. *Mean LSAT, UGPA, and FYA differences within school by gender subgroup*

Differential Validity

The relationships between the predictor variables (LSAT score and UGPA) and the criterion variable (FYA) are measured through the computation of correlation coefficients. Correlation coefficients can range in value from –1 to 1, where 1 represents a perfect positive linear relationship. The correlation coefficients for each gender subgroup were calculated separately by law school and averaged across all schools.

Table 6 provides the correlation coefficients between the predictors (LSAT score and UGPA)—both alone and in combination—and the FYA. For both male and female students, the combination of LSAT score and UGPA results in higher correlations than either individual predictor alone. Results also show that for both gender subgroups, LSAT score has a stronger correlation with FYA than UGPA does.

To further evaluate the results presented in Table 6, correlation differences between the male and female subgroups were calculated using Fisher's *r*-to-*z* transformation (Fisher, 1921) to assess whether the differences were statistically significant. None of the calculated differences were found to be significantly different from one another at a significance level of $p < .05$.

TABLE 6
Correlation of LSAT score and UGPA with FYA by gender subgroup

Gender Subgroup	No. of Schools	No. of Students	Correlation		
			LSAT *M* (*SD*)	UGPA *M* (*SD*)	LSAT & UGPA *M* (*SD*)
Male	158	51,901	0.36 (0.08)	0.24 (0.09)	0.45 (0.08)
Female	158	47,124	0.44 (0.09)	0.29 (0.11)	0.53 (0.08)

Differential Prediction

Whereas the above evaluation of the correlation coefficients provides a measure of overall bias across the range of the predictor, the evaluation of differential prediction examines whether each predictor or combination of predictors systematically over- or underpredicts FYA for gender subgroups. Differential prediction was quantified by subtracting the observed FYA from the predicted FYA for each model and comparing the average of this difference based on gender subgroups. Negative average values indicate underprediction, whereas positive average values indicate overprediction.

Using least-squares regression, separate equations were derived to predict law school FYA for the total group of law school students within each individual law school for LSAT score alone, UGPA alone, and the combination of both predictor variables. The weighted averages of the mean residuals between observed FYA and predicted FYA for each prediction equation/subgroup combination are provided in Table 7.

TABLE 7
Summary of mean residuals between predicted and observed FYA by gender subgroup

Gender Subgroup	LSAT Score	UGPA	LSAT & UGPA
Male	0.28	–0.61	–0.15
Female	–0.31	0.67	0.16

Table 7 reveals that when FYA is estimated from a regression equation based on data for all students, LSAT score tends to slightly underpredict the performance of female law students (residual of –0.31) and slightly overpredict the law school performance of male law students (residual of 0.28). Conversely, the use of UGPA alone tends to slightly overpredict the performance of female law students (0.67) and slightly underpredict the performance of male law students (–0.61). While the over- and underprediction for both of these variables in isolation is quite small, the magnitude when using UGPA alone is approximately double that observed for LSAT score alone. Lastly, results indicate that the prediction equation using a combination of both LSAT score and UGPA led to the most accurate prediction of FYA for both gender subgroups, with a residual of –0.15 for male students and a residual of 0.16 for female students. While the direction of the residuals presented in Table 7 may be interpreted to indicate either over- or underprediction, it is worth noting that the magnitude of the difference from zero for all reported residuals is small enough to be considered inconsequential.

Conclusions

The purpose of this study was to address questions of differential validity and differential prediction for gender subgroups. Validity coefficients between predictors (LSAT score and UGPA) and FYA were evaluated for male and female law students, and regression equations were fit to the data using LSAT score, UGPA, and the combination of LSAT score and UGPA as predictors of FYA. Differences between predicted and observed FYA for each model were evaluated to assess whether the use of a common regression equation could systematically exclude members of gender subgroups in the admission process.

Results indicate that the use of UGPA as a sole predictor seemed to produce the greatest amount of differential prediction between male and female law students. In fact, the magnitude of over- and underprediction based on UGPA alone was approximately twice that observed for LSAT score alone. Still, the magnitude of over- and underprediction observed for UGPA alone was quite small. Results also indicate that the use of the combination of both LSAT score and UGPA in the prediction equation led

to the least amount of differential prediction for both male and female law students. In fact, the over- and underprediction observed with both predictors included in the model was quite small for both gender subgroups.

At least two caveats should be remembered when evaluating the results of this study. First, only differences in average predicted performance were analyzed. That is, the performance of individuals within a subgroup whose FYAs are overpredicted on average may still be underpredicted, and vice versa. Second, differential prediction is only one aspect of an overall construct validity evaluation. Other aspects of validity should also be considered when deciding whether the use of any test scores is valid.

References

American Bar Association. (2016). Statistics [Data files]. Retrieved from http://www.americanbar.org/groups/legal_education/resources/statistics.html.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Burton, N. W., & Wang, M. (2005). *Predicting long-term success in graduate school: A collaborative validity study* (Graduate Records Examination Board Research Report No. 99-14R and Educational Testing Service, RR-05-03). Princeton, NJ: Educational Testing Service.

Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron, 1,* 3–32.

Linn, R. L. (1978). Single-group validity, differential validity, and differential prediction. *Journal of Applied Psychology, 63,* 507–512.

Mattern, K. D., Patterson, B. F., & Kobrin, J. L. (2012). *The validity of SAT® scores in predicting first-year mathematics and English grades* (College Board Research Report 2012-1). New York, NY: The College Board.

Noble, J. (2003). *The effects of using ACT composite score and high school average on college admission decisions for racial/ethnic subgroups* (ACT Research Report 2003-01). Iowa City, IA: American College Testing.

Norton, L., Suto, D., & Reese, L. (2013). *Analysis of differential prediction of law school performance by gender subgroups based on 2008–2010 entering law school classes* (LSAC Technical Report, TR 13-01). Newtown, PA: Law School Admission Council.

Shaw, E. J., Kobrin, J. L., Patterson, B. F., & Mattern, K. D. (2012). *The validity of the SAT® for predicting cumulative grade point average by college major* (College Board Research Report 2012-6). New York, NY: The College Board.

Suto, D., Norton, L., & Reese, L. (2007). *Analysis of differential prediction of law school performance by gender subgroups based on 2002–2004 entering law school classes* (LSAC Technical Report, TR 07-01). Newtown, PA: Law School Admission Council.

Suto, D., Norton, L., & Reese, L. (2010). *Analysis of differential prediction of law school performance by gender subgroups based on 2005–2007 entering law school classes* (LSAC Technical Report, TR 10-01). Newtown, PA: Law School Admission Council.