

Why and How to Reform Or Abolish the LSAT:

*Why the proposed changes to ABA Standard 503 accomplish
nothing and how that can be remedied*

Includes statistical and policy support for eliminating inequities in the LSAT
test-taker's preparation process and law school admissions

An Urgent Communiqué from an LSAT Survivor
and Prospective Law Student

Completed March 24, 2011

Revised April 3, 2011

Al Alston

alalston@thenile.com

215-769-3333



3117 Master Street Philadelphia, PA 19121

Phone: 215-769-3333 Fax: 215-769-3980

www.songhaicity.org

Preface

First, this is a responsible, humble attempt to influence the thoughts of policy makers at the American Bar Association (ABA), other accreditation bodies, law schools and any other organizations that impact or are impacted by the LSAT, LSAC, and law school admissions practices. This communiqué represents an effort to use what I know about statistics, the law, the social contract and fairness to expose what I consider to be an extraordinary injustice. Namely, I have concluded that LSAC is engaged in a brilliant marketing scheme that heralds unconscionable, poorly performing, pseudo-scientific metrics as legitimate and unbiased, when they almost certainly know that nothing could be further from the truth. My goal is to help expose the fraud being perpetrated on 150,000 test-takers each year and to offer support for policymakers who recognize that it is well past the time to eliminate or dramatically modify the LSAT.

The November 2010 proposed revisions to ABA Standard 503 do not begin to address the underlying issues. I have incorporated several suggestions about how to modify the proposed changes so that they remove the ABA's tacit support for scientifically invalid tests, protect prospective law students from an intolerable violation, and help to shield law schools from continual litigation.

Second, though the result of significant time and research, this cannot be considered a scholarly paper and I don't make any claims to that effect. I don't have a Ph.D. in statistics or education research or anything else. My undergraduate background was in electrical engineering, probability analysis, econometrics, public policy and sociology. My graduate work was in electrical engineering, with courses in stochastic processes. While I have forgotten more from those fields than I recall, my review of many aspects of those disciplines during the preparation of this document provided me with sufficient insight to complete what I believe is a reasonable and thought-provoking document.

I simply call this document a communiqué, for lack of a better term, to avoid any misconceptions about whether I believe I speak with any more authority than the above description of my skills and insights suggests. Accordingly, please be prepared to see the word "I" frequently. I will express opinions based on research or personal experience, but I never masquerade them as facts and always provide a basis for them. My research of the literature was not exhaustive – time, money and, frankly, lack of exposure in some subject areas limited what I could review (and for how long, in some cases).

The scope of this communiqué is significant. There are likely to be errors present. I apologize, in advance, for any errors I made (hopefully, few) or attributions that I neglected (hopefully, none). I will correct anything brought to my attention in a future draft/version, if there is one.

Table of Contents

1. EXECUTIVE SUMMARY	1
1.1. THE LSAT IS INVALID.....	1
1.2. THE LSAT IS INEQUITABLE	1
1.3. RECOMMENDATION FOR THE ABA STANDARDS COMMITTEE.....	1
1.3.1. <i>Revision of Standard 503 to prohibit LSAT use.....</i>	<i>1</i>
1.3.2. <i>If Standard 503 is modified to allow each law school to determine LSAT use</i>	<i>2</i>
1.4. RECOMMENDATIONS FOR LSAC – MODIFYING THE LSAT	2
2. DISCUSSION, SUMMARY AND RECOMMENDATIONS.....	3
2.1. INTRODUCTION	3
2.1.1. <i>LSAT validity.....</i>	<i>3</i>
2.1.2. <i>LSAT test-taker preparation inequities</i>	<i>3</i>
2.1.3. <i>Responsibilities of the ABA and LSAC.....</i>	<i>4</i>
2.2. FINDINGS	4
2.3. RECOMMENDATIONS.....	5
2.3.1. <i>Recommendations for the ABA Standards Committee</i>	<i>5</i>
2.3.2. <i>Recommendations for all policy makers whether within LSAC, the ABA, law schools, Congress, etc.....</i>	<i>7</i>
3. INTRODUCTION	10
3.1. THE LSAT AND LAW SCHOOL ACCREDITATION	10
3.2. MUST DECOUPLE THE LSAT FROM THE LAW SCHOOL ADMISSIONS PROCESS	11
3.2.1. <i>Existing Standards.....</i>	<i>11</i>
3.2.2. <i>Proposed Standards.....</i>	<i>12</i>
3.3. THE URGENCY OF NOW.....	12
3.4. ORGANIZATION OF REMAINING SECTIONS - WHAT WE WILL SHOW	13
3.4.1. <i>Proposal A. Modify the LSAT and its Test Administration Procedures to Reduce or Eliminate Inequities.....</i>	<i>13</i>
3.4.2. <i>Proposal B. LSAC’s studies contain errors, peculiar choices, spin that distorts the poor performance of the LSAT, and a clear refusal to investigate questions damaging to the LSAT.....</i>	<i>14</i>
3.4.3. <i>Proposal C. LSAT’s predictability rates of FYA justify rescinding requirements that students take the LSAT and recommend a 100% compensatory model.....</i>	<i>14</i>
3.4.4. <i>Proposal D. Racial and ethnic LSAT disparities must be eliminated within the next twenty years or diversity efforts may be ruled unconstitutional.....</i>	<i>15</i>
3.4.5. <i>Proposal E. Ways to reduce LSAT inequities, other than modification or abolition.....</i>	<i>15</i>
3.4.6. <i>Conclusion</i>	<i>15</i>
3.4.7. <i>Exhibit A</i>	<i>15</i>
4. PROPOSAL A. MODIFY THE LSAT AND ITS TEST ADMINISTRATION PROCEDURES TO REDUCE OR ELIMINATE INEQUITIES.	17
4.1. TODAY’S LSAT.....	17
4.2. WHY TEST-TAKER PREPARATION IS NECESSARY FOR THE LSAT.....	17
4.3. THE LSAT TEST-TAKER’S PREPARATION BURDEN	18
4.3.1. <i>Required: Time.....</i>	<i>18</i>
4.3.2. <i>Required: Money</i>	<i>19</i>
4.3.3. <i>Required: Personal/Professional sacrifices</i>	<i>20</i>
4.4. INEQUITIES RESULTING FROM THE LSAT TEST-TAKER’S PREPARATION BURDEN	21
4.4.1. <i>Pervasiveness and intensity of inequities from LSAT test-taker’s preparation burden.....</i>	<i>21</i>
4.5. THE LSAT TEST-TAKER’S PREPARATION BURDEN AND LSAT SCORES.....	24
4.5.1. <i>LSAC Research: LSAT scores decline with age – implications for preparedness</i>	<i>24</i>
4.5.2. <i>LSAC Research: LSAT Repeater Data.....</i>	<i>25</i>

4.5.3.	<i>Anecdotal evidence</i>	27
4.5.3.1.	Repeat test-takers.....	27
4.5.3.2.	Personal account.....	27
4.6.	MODIFYING THE LSAT.....	28
4.6.1.	<i>Evidence, assumptions and action</i>	28
4.6.2.	<i>Assumptions and constraints</i>	29
4.6.3.	<i>Scientific research</i>	29
4.6.3.1.	Ability is independent of speed, extra time tends to increase scores.....	29
4.6.4.	<i>LSAC's analysis of "speededness" in the LSAT</i>	32
4.6.5.	<i>Applicability of Research to the LSAT</i>	34
4.7.	RECOMMENDED LSAT MODIFICATIONS.....	35
4.7.1.	<i>Modification Set #1: To minimize test preparation inequities</i>	36
4.7.1.1.	1(a) Employ untimed or extra-time testing universally.....	36
4.7.1.2.	1(b) LSAC will provide comprehensive disclosures.....	36
4.7.1.3.	1(c) Eliminate questions and question types that do not add value.....	37
4.7.1.3.1.	LSAC research demonstrates that AR and RC sections add redundancies.....	37
4.7.1.3.2.	Continued inclusion of AR and RC wastes resources.....	39
4.7.1.3.3.	Removal of AR and RC would decrease the intense anxiety associated with LSAT.....	39
4.7.1.3.4.	Removal of AR and RC would decrease the inequities in student preparation for the LSAT.....	40
4.7.1.3.5.	Removal of AR and RC would decrease racial, ethnic and gender biases present in the LSAT.....	40
4.7.2.	<i>Modification Set #2: To consistently measure ability, introduce a power score</i>	41
4.7.2.1.	The Mastery score, what it would assess, how it would be used.....	41
4.7.3.	<i>Modification Set #3: To enhance differentiation among the best prepared, introduce a speed score</i>	41
4.7.3.1.	The Intensity score, what it measures, how it would be used.....	42
4.7.3.2.	Law school application of Intensity scores.....	42
4.7.4.	<i>Combining both Mastery and Intensity scores</i>	43
4.7.4.1.	Additional data points can be gathered.....	43
4.7.4.2.	Avoiding the stigmatization of Mastery scores.....	43
4.7.4.3.	Importance of defense measures that would lessen or eliminate opportunities to exploit the new regime.....	44
4.7.4.4.	Mastery scores and LR-only testing need each other.....	45
4.7.4.5.	Assumptions must be validated with independent studies.....	45
5.	PROPOSAL B. LSAC'S STUDIES CONTAIN ERRORS, PECULIAR CHOICES, SPIN THAT DISTORTS THE POOR PERFORMANCE OF THE LSAT, AND A CLEAR REFUSAL TO INVESTIGATE QUESTIONS DAMAGING TO THE LSAT.	46
5.1.	ERRORS AND PECULIAR CHOICES IN STUDIES.....	46
5.2.	SPIN THAT DISTORTS THE POOR PERFORMANCE OF THE LSAT.....	46
5.2.1.	<i>The pervasive, deceptive and inappropriate use of the correlation coefficient (r) and refusal to use the coefficient of determination (R²)</i>	46
5.2.1.1.	Introduction.....	46
5.2.1.2.	Correlation coefficient (r) as "proportion of variance" fails as an intuitive measure.....	47
5.2.1.3.	The coefficient of determination, R ² , a standard for evaluating correlation.....	47
5.2.1.4.	Correlation coefficient (r) is misleading and inferior to coefficient of determination (R ²).....	49
5.2.1.5.	Gauging standards for correlation and the state of behavioral sciences.....	50
5.2.1.6.	LSAT's historical data revisited, using coefficients of determination and alienation.....	51
5.2.1.6.1.	How much of FYA is NOT explained by LSAT – National Correlation Studies.....	52
5.2.1.6.2.	How much of FYA is NOT explained by LSAT – Differential Prediction for Racial/Ethnic Subgroups.....	53
5.2.1.6.3.	How much of FYA is NOT explained by LSAT – Differential Prediction for Gender Subgroups.....	54
5.2.1.6.4.	How much of FYA is NOT explained by LSAT – Item Type Validity.....	55
5.2.1.6.5.	Observations about LSAT's historical performance as a measure of FYA.....	55
5.2.1.6.6.	The psychology vs. responsibility of declaring LSAT's results poor and unacceptable.....	56
5.2.2.	<i>LSAC researchers repeatedly draw unjustified conclusions using correlation coefficient (r)</i>	57
5.2.2.1.	Predictive Validity of the LSAT.....	57

5.2.2.2.	Analysis of Differential Prediction of Law School Performance - Racial/Ethnic Subgroups ..	58
5.2.2.3.	Conclusions on LSAC’s interpretation of correlation coefficient.....	58
5.3.	CLEAR REFUSAL TO INVESTIGATE QUESTIONS DAMAGING TO THE LSAT	59
5.3.1.	<i>Missing analysis of how LSAT preparation impacts LSAT performance</i>	<i>59</i>
5.3.1.1.	LSAT scores decline with age – reasons never explored by LSAC.....	60
5.3.1.2.	LSAT scores decline with age – implications for preparedness	60
5.3.2.	<i>Use of mean when median and percentiles would be more informative.....</i>	<i>60</i>
5.3.2.1.	Misleading analysis of repeat test-takers	61
5.3.2.1.1.	Performance study of 2000 – 2007 repeaters	61
5.3.2.1.2.	Information hidden in RepeaterData.pdf.....	62
5.3.2.1.3.	Validity of Law School Admission Test Scores for Repeat Test Takers.....	63
5.3.2.1.4.	Dramatic differences in average LSAT scores raise questions about utility of LSAT.....	64
6.	PROPOSAL C. LSAT’S PREDICTABILITY RATES OF FYA JUSTIFY RESCINDING REQUIREMENTS THAT STUDENTS TAKE THE LSAT AND RECOMMEND A 100% COMPENSATORY MODEL.....	65
6.1.	LSAT’S POOR PERFORMANCE CANNOT JUSTIFY ACCREDITATION MANDATE FOR LSAT.....	65
6.2.	PRECAUTIONS ABOUT WITHDRAWING STANDARD 503 AND ALTERING STANDARD 501	65
6.3.	TOWARDS A COMPENSATORY MODEL FOR EVERYONE	66
7.	PROPOSAL D. RACIAL AND ETHNIC LSAT DISPARITIES MUST BE ELIMINATED WITHIN THE NEXT TWENTY YEARS OR DIVERSITY EFFORTS MAY BE RULED UNCONSTITUTIONAL.	67
7.1.	ARGUMENT	67
7.2.	DIVERSITY IS COMPELLING, BUT RACE/ETHNICITY LSAT CORRECTIONS WILL BE UNCONSTITUTIONAL	67
7.3.	RACIAL/ETHNIC LSAT DISPARITIES ALMOST IMPERVIOUS TO CHANGE.....	69
7.4.	LSAT MODIFICATIONS OR REPLACEMENTS MUST ADDRESS PREPARATION INEQUITIES AND LAW SCHOOL OBJECTIVES	71
8.	PROPOSAL E. WAYS TO REDUCE LSAT INEQUITIES, OTHER THAN MODIFICATION OR ABOLITION OF THE LSAT.....	72
8.1.	COMMENCE INDEPENDENT STUDIES	72
8.2.	IMMEDIATE ACTIONS.....	72
8.2.1.	<i>Create free, high-quality LSAT preparation courses at university level.....</i>	<i>73</i>
8.2.2.	<i>Make truthful, continuous and pervasive public service announcements about LSAT.....</i>	<i>73</i>
9.	CONCLUSIONS.....	75
10.	EXHIBIT A: ERRORS AND PECULIAR CHOICES IN LSAT STUDIES	78
10.1.	INTRODUCTION	78
10.2.	A 30-SECOND (VISUAL) PRIMER	78
10.3.	ANALYSIS OF DIFFERENTIAL PREDICTION STUDIES – RACE/ETHNICITY	79
10.4.	PREDICTIVE VALIDITY OF THE LSAT - NATIONAL.....	82
10.5.	LSAT ITEM-TYPE VALIDITY	83
10.6.	GENERAL	84

Tables

TABLE 1: POPULATIONS LIKELY IMPACTED BY EXCESSIVE LSAT TEST-TAKER’S PREPARATION BURDEN, 2008 – 10.....	23
TABLE 2: IMPROVEMENTS IN LSAT SCORES AMONG REPEAT TEST-TAKERS.....	25
TABLE 3: NUMBER OF ADDITIONAL CORRECT ANSWERS FOR EACH IMPROVED SCORE RANGE.....	26
TABLE 4: GENERAL RECOMMENDATIONS FOR REFORMING LSAT, BASED ON ASSUMPTIONS & CONSTRAINTS.....	35
TABLE 5: CORRELATIONS BETWEEN LSAT SECTIONS AND FYA.....	38
TABLE 6: PERCENT ANSWERED CORRECTLY FOR EACH SECTION, BY RACE, ETHNICITY & GENDER.....	40
TABLE 7: HOW MUCH OF FYA IS NOT EXPLAINED BY LSAT – CORRELATION STUDIES FROM 1995 – 2008.....	52
TABLE 8: HOW MUCH OF FYA IS NOT EXPLAINED BY LSAT – RACE/ETHNICITY 2005 - 2007.....	53
TABLE 9: HOW MUCH OF FYA IS NOT EXPLAINED BY LSAT – GENDER 2005 - 2007.....	54
TABLE 10: HOW MUCH OF FYA IS NOT EXPLAINED BY LSAT – ITEM TYPE 1993 - 1996.....	55
TABLE 11: MEAN LSAT SCORES BY RACE/ETHNICITY – 1993 - 1994.....	70
TABLE 12: MEAN LSAT SCORES BY RACE/ETHNICITY – 2009 - 2010.....	70

1. EXECUTIVE SUMMARY

The Law School Admission Council (LSAC) is a nonprofit corporation that administers the LSAT, a standardized test that virtually every law school in the United States requires its applicants to take, as a result of American Bar Association (ABA) Standard 503. For decades, LSAC has created “research” reports and issued public statements that portray the LSAT as an objective, valid scientific tool. However, the facts support a different interpretation, that the LSAT is an inequitable failure as a measure of ability. Therefore, either the ABA must modify Section 503 to abolish the LSAT mandate *for all law schools* or it must support major changes to the LSAT (and LSAC must implement them) immediately. The proposed removal of Standard 503 and slight modifications to Standard 501 do neither, but their existence creates an opportunity for substantial progress. In this communiqué, written by an LSAT test-taker who experienced many of the issues described herein, the above arguments are fully supported and detailed recommendations provided for addressing the problems identified.

1.1. The LSAT is invalid

A careful examination of LSAC’s own reports reveals that the LSAT is a statistically, methodologically and inherently invalid measure of ability and predictor of law school performance that can no longer be justified. Specifically, LSAC’s reports show that:

- The LSAT is paced too quickly to measure ability consistently, so that a significant component of what it actually measures is speed (rate of question completion).
- The LSAT is a poor predictor of law school performance, generally failing to explain approximately 84% - 95% of the variation in first-year law school performance.
- Over the last ten years, the predictive capabilities of the LSAT have dropped precipitously – 32% (using R^2) and 17% (using r).

1.2. The LSAT is inequitable

LSAC deceptively asserts that “some preparation” is required of those taking the LSAT. Conventional wisdom, however, holds that three months of preparation, as well as the purchase of potentially expensive materials or registration at costly commercial test preparation centers is required. The requisite time and money associated with preparation are potentially massive barriers that disproportionately impact low-income applicants and those with considerable personal or business responsibilities. These inequities might also explain the poor performance of older applicants, underrepresented groups, and, to a lesser extent, females. The LSAT, therefore, is invalid, inequitable and its continued use creates unacceptable and, possibly, illegal admissions decisions.

1.3. Recommendation for the ABA Standards Committee

1.3.1. Revision of Standard 503 to prohibit LSAT use

The most obvious means of addressing the issues posed by the LSAT is for the ABA to remove the root of problem, by modifying Standard 503 to prohibit any law school from basing an admissions decision on the LSAT. The Nov. 2010 Standards draft, removing Section 503 and altering 501, does not accomplish this. However, the following might:

<p><u>Suggested modification to Standard 501(a):</u> A law school shall maintain sound admission policies and practices, consistent with its mission and the objectives of its educational program, and which does not include an admissions exam of any sort.</p>

Today's special-case compensatory admission models can then be used in all cases, as the proposed Interpretation 501-1 approaches. This interpretation could be expanded, however, to identify more than just "obstacles overcome" for diversity maintenance.

1.3.2. If Standard 503 is modified to allow each law school to determine LSAT use
The November 2010 Standards draft, removing Standard 503, and revising 501, would allow individual law schools to determine for themselves whether they will require applicants to take the LSAT or other standardized tests. Unfortunately, this would do absolutely nothing to address the invalidity or inequities of the LSAT. We must note, as a contingency, however, that such changes would still bring responsibilities for the ABA.

Suggested, conditional modifications to revised Interpretation 501-1: Sound admissions policies and practices may include consideration of ~~admission test scores~~, undergraduate course of study and grade point average, extracurricular activities, work experience, performance in other graduate or professional programs, performance in a law school pre-admission program, relevant demonstrated skills, and obstacles overcome.

Consideration of admission test scores is permissible at a given law school only after there has been a finding by the ABA or the Dean of that law school that the test is scientifically valid, does not impose inequitable hardships on the applicants, and measures only ability essential for law school success. This finding must be recertified each academic year.

Suggested, conditional new Interpretation 501-4 (or 5): When a law school's sound admissions policy includes an admission test, that school must:

- Accept only test results from an untimed or extra-time test administration.
- Provide and publish monthly a clear, accurate and complete statement about what the chosen standardized test measures, what it does not measure, how long it takes to prepare for the exam, how one should prepare for the exam.
- Provide a high-quality test-preparation course free of charge for all applicants or pay the cost of a high-quality commercial one.
- Ensure that the test provider offers a free catalog of all question types and how to solve each.
- Work with the test provider to ensure the exam's continual validity and equity.

1.4. Recommendations for LSAC – Modifying the LSAT

If the LSAT is retained as a requirement for any or all law schools, LSAC must support the recommendations above, because they will **help** to make the LSAT valid and more equitable. However, modifications to the LSAT would **still be required** to ensure the test's validity and fairness for all test-takers, while helping schools maintain diversity and retain measures for the most selective universities. These modifications include steps to:

- Remove the Reading Comprehension (RC) and Analytical Reasoning (AR) sections, which do not add significantly to the exam's predictive value, but increase prep time.
- Create a power score (% correct of questions attempted) – for proficiency, diversity.
- Create a speed score (# answered correctly) and infrastructure – for selectivity.

For all interested parties, there must be additional and immediate analysis of LSAC's studies, the critical questions LSAC never explored, and the above recommendations.

2. DISCUSSION, SUMMARY AND RECOMMENDATIONS

Below, we introduce the fundamental challenge of assessing the validity of the LSAT and identifying the test-taker preparation issues associated with the exam. We discuss the findings of our six weeks of research into each. Finally, we summarize our two sets of recommendations for addressing what we learned: those for the ABA Standards Committee and those for everyone else, including the remainder of the ABA, law schools, prospective law students, and every other interested party.

2.1. Introduction

This communiqué addresses two distinct issues arising from LSAT use: whether it is valid as a test of ability and whether the required preparation for it can be considered fair to all test-takers.

2.1.1. LSAT validity

In assessing validity, we examine the academic literature on standardized testing, the use of statistical analysis in behavioral sciences, and LSAC's own research to identify acceptable standards as well as the metrics chosen for the LSAT. The term "validity" arises frequently when assessing whether a standardized test is timed properly so that the great majority of test-takers have the opportunity to complete the exam. We explore this mode of analysis in detail. LSAC conducts regular statistical comparisons between LSAT scores and first-year law school GPA's ("FYA's") to determine how well LSAT scores can serve as predictors. We examine LSAC's methodology, results, analysis and conclusions and submit our findings as a partial answer to whether the LSAT is a valid measure of ability. *We do not examine whether specific question types actually measure what LSAC indicates they measure or whether they are, otherwise, appropriate.*

2.1.2. LSAT test-taker preparation inequities

This is an area that is rarely discussed in academic literature, but it is an essential topic of discussion for prospective law school students. We demonstrate that it is fundamental to understanding the impact of the LSAT on the lives of every law school applicant and whether the LSAT is a meaningful or helpful gauge of their abilities.

LSAC states that the LSAT is designed to measure skills "essential" to success in law school and that "some preparation" will be necessary by most. However, this is a gross understatement of the extraordinary preparation required by the great majority of prospective students. Conventional wisdom says that the typical preparation time is three months. Policymakers must face a central, perhaps rhetorical question: whether every law school applicant has (a) three months available to study for the LSAT, (b) the financial resources to fund that study, and (c) sufficient awareness of the specific skills that must be mastered in any preparation process.

When one invariably answers no to this question, then the impacts of preparation inequities must be measured and addressed. While additional study would be helpful, we believe that there is already sufficient evidence available to describe the inequities, those who are likely to be impacted, and what we can do to reduce or eliminate this solemn, pernicious and pervasive injustice. Universities have been addressing the outcomes of

this corrupted system by fashioning diversity programs, many of which will soon be unconstitutional per *Grutter v. Bollinger*, 539 U.S. 306 (2003).

However, even these programs might be failing to identify qualified students because they compensate for LSAT scores, not the reasons for those scores: the LSAT preparation process and the LSAT itself.

2.1.3. Responsibilities of the ABA and LSAC

There is a narrow window of time remaining to create a more equitable admissions process that will be reflective of the great diversity of the nation, that will be safely within the bounds of the law, that will reduce LSAT anxieties, and that will allow everyone qualified to be an attorney to compete fairly for the open seats at law schools. But, no matter how wide or narrow that window, three things are clear. The ABA must revise its standards to address the failures of the LSAT; the ABA must take a more proactive and visible role in shaping the course of admissions process for law schools; and LSAC must address the criticisms of the LSAT without delay.

2.2. Findings

There is a cottage industry of LSAT test preparation materials and courses, most of which have price tags that make them inaccessible to the poor, large cross-sections of older applicants, underrepresented people of color, and those with significant family responsibilities. Many in those same groups, as well as those with considerable career or business responsibilities, can be expected to have commitments that would preclude them from devoting significant time to LSAT preparation.

LSAC, unfortunately, has refused to measure, quantitatively, how test preparation affects test scores and how demographic characteristics affect preparation. According to their cryptic reports, however, nearly 40% of repeat test-takers improve their scores moderately or significantly upon their second or third attempt. This is an indication that preparation is a strong factor in LSAT performance. LSAC's research also demonstrates that older prospective students of all races and ethnicities, as well as underrepresented people of color perform far worse than young White or Asian test-takers. These are the same groups that one would expect to be disproportionately impacted by inequities arising from requisite, extensive preparation for the LSAT.

LSAC's own research also demonstrates that parts of the LSAT are duplicative; that the LSAT is a poor indicator of law school performance; that its predictive abilities have *worsened* precipitously over the last ten years. The reliance on the correlation coefficient (r) and not the more appropriate and intuitive coefficient of determination (R^2) to measure validity is almost a fetish with LSAC and among its supports. However, we rely upon the insights of a noted statistician to explore the psychology of this adherence and suggest a resetting of perspective, where appropriate.

LSAC's research also demonstrates that LSAT exams are paced too quickly to measure ability consistently. This is a statement that the exam is "speeded" and is a direct challenge to the validity of the LSAT, as it would be to virtually any standardized test.

LSAC's researchers argue that LSAC has greatly underestimated the speededness issues associated with the LSAT. The current evidence, therefore, points to one conclusion: the LSAT does not measure ability consistently; it measures, to an unacceptable degree, speed. And, it is precisely *speed* that can be most improved during a three-month preparation period.

LSAC has engaged in a pattern of deception to hide these facts in plain sight. Their studies stop short of uncovering what the LSAT really measures, simply because LSAC has refused to examine questions about preparation strategies, test-taker hardships, actual test completion rates (factoring in rapid-guessing), and associated outcomes.

Combining test preparation inequities with a fatally flawed test suggests that the LSAT must be abolished through a revised Standard 503 or both the LSAT and how it is administered must be significantly modified. As well, far more attention must be given to disclosures about the LSAT, the specific skills it attempts to measure, test-taking strategies, and the involvement of universities in the formal test preparation process. In the end, however, accrediting agencies and law school admissions offices must re-examine their support for the LSAT and create far less problematic evaluative and selection procedures.

The November 2010 Standards draft, which removes Section 503 and slightly modifies Standard 501 and its Interpretations, does not address any of the above, with the possible exception of Interpretation 501-1, which is a starting point for a list of "sound" admission criteria.

2.3. Recommendations

To address LSAT inequities and poor reliability, and to arrive at admissions policies that promote diversity, do not stigmatize, and which obviate the need to factor in the poorly measured low scores of underrepresented people of color, we propose, and request that you consider, the following recommendations in your accrediting and/or oversight role. Note that some are applicable should the LSAT be abolished as a law school entrance requirement; others presume that the LSAT requirement will remain in some form; others are applicable in either case. As well, some are addressed only to the ABA Standards Committee.

2.3.1. Recommendations for the ABA Standards Committee

A. Regarding abolishment of the LSAT requirement in Standard 503:

1. Modify Standard 503 of the *2010-2011 ABA Standards for Approval of Law Schools* so that it prohibits law schools from requiring prospective students to take the LSAT. Until there is a valid and reliable replacement, no standardized tests should be required by any law school. The November 2010 removal of Standard 503 and the alterations to Standard 501-1 do not accomplish this and are not improvements over Standard 503. Abolition language could be:

Suggested modification to Standard 501(a): A law school shall maintain sound admission policies and practices, consistent with its mission and the objectives of its educational program, **and which does not include an admissions exam of any sort.**

2. Alternatively, modify Standard 503 of the 2010-2011 ABA Standards for Approval of Law Schools so that it abolishes the ABA mandate for standardized tests, but imposes tough certification measures for any tests that law schools wish to authorize and require. As noted above, the removal of Standard 503 does not address the concerns regarding the LSAT and the proposed modifications to Interpretation 501-1 leave open avenues for abuse. Interpretation 501-1 allows each law school to determine whether it will require its applicants to take the LSAT or any other standardized test, but it does not address the inequities and the invalid nature of the LSAT. In fact, it forces those seeking to improve the fairness of admissions processes, while increasing or maintaining diversity, to fight 200 separate battles. This can be cured through changes to the proposed Interpretation 501-1, however, as demonstrated below.

Suggested, conditional modifications to revised Interpretation 501-1: Sound admissions policies and practices may include consideration of ~~admission test scores~~, undergraduate course of study and grade point average, extracurricular activities, work experience, performance in other graduate or professional programs, performance in a law school pre-admission program, relevant demonstrated skills, and obstacles overcome.

Consideration of admission test scores is permissible at a given law school only after there has been a finding by the ABA or the Dean of that law school that the test is scientifically valid, does not impose inequitable hardships on the applicants, and measures only ability essential for law school success. This finding must be recertified each academic year.

3. In lieu of a required LSAT, mandate that the law schools you accredit or advise adopt a compensatory admission model 100% of the time, where any LSAT scores or those from any other standardized test are considered unreliable and not factored into the admissions decision. In new compensatory admission models, there must be a diminished reliance upon race, ethnicity and gender. Diversity might be sought through constitutionally acceptable factors correlated with the disadvantages suffered by underrepresented people of color and women. The use of “obstacles overcome” in proposed Interpretation 501-1 goes to this, though additional guidance could be offered. Weighting high school class standing might help to achieve this, as per the Texas Top 10% Rule. However, more study is required.

B. Assuming an LSAT requirement will be permitted by some version of Standard 503 or the standard’s removal:

4. If law schools are permitted to retain the LSAT as a requirement, create a new standard which requires that the LSAT or any other standardized test be administered on an untimed or extra-time basis (approximately 30 additional minutes per section). This would reduce or eliminate the speeded nature of the test and better allow it to measure ability, not speed. This would also minimize test-preparation advantages because intensive drilling to improve speed would no longer be necessary. The LSAT has question types in reasoning and reading comprehension that can be solved with additional time, but only if the prospective student has the ability to do so. (This is unlike question types such as vocabulary or word associations, which a student either knows or does not know.) The College Board (owner of the SAT), Pearson

Education (owner of the Stanford Achievement Test Series) and other educational testing entities are already giving standardized tests in untimed and extra-time settings, so this is a practical modification.

5. If law schools are permitted to retain the LSAT as a requirement, create a new standard which **encourages** pre-law programs to offer high-quality, full credit LSAT preparation courses and **mandates** that law schools which rely upon the LSAT for admissions decisions offer free, high-quality courses. Perhaps grades from the law school LSAT preparation courses could be substituted for the LSAT scores or a modified LSAT could be the final. The undergraduate option eliminates cost and time from the equation of LSAT preparedness. The law school option eliminates cost and perhaps any over-reliance on the LSAT itself.
6. If law schools are permitted to retain the LSAT as a requirement, create a new standard which requires LSAC to provide, free of charge, full disclosure of the following: what the LSAT attempts to measure and how well it succeeds or fails at that mission; comprehensive descriptions of every type and variant of question that could appear on the LSAT and how to solve each and every one of them; comprehensive test-taking strategies; how much time and financial commitment are required to perform maximally.

The above three recommendations can be summarized by changes to the revised Standard 501, as follows.

Suggested, conditional new Interpretation 501-4 (or 5): When a law school's sound admissions policy includes an admission test, that school must:

- Accept only test results from an untimed or extra-time administration.
- Provide and publish monthly a clear, accurate and complete statement about what the chosen standardized test measures, what it does not measure, how long it takes to prepare for the exam, how one should prepare for the exam.
- Provide a high-quality test-preparation course free of charge for all applicants or pay the cost of a high-quality commercial one.
- Ensure test provider offers a free catalog of all question types and how to solve each.
- Work with the test provider to ensure the exam's continual validity and equity.

2.3.2. Recommendations for all policy makers whether within LSAC, the ABA, law schools, Congress, etc.

A. *Regarding modifying the LSAT:*

7. If law schools are permitted to retain the LSAT as a requirement, encourage law schools to consider the scores from only the two Logical Reasoning (LR) sections. LSAC's own data demonstrate that LR scores are as good (or bad) predictors of law school success as those of the entire LSAT exam. Encourage LSAC to stop administering the Analytical Reasoning (AR) and Reading Comprehension (RC) sections. Eliminating AR and RC would better focus a student's preparation efforts, lower the preparation costs and lessen racial differentiation in LSAT scores, without harming the LSAT's (limited) utility. An LR-only test, with extra-time, would be

completed about as quickly as today's.

8. If the LSAT is moved to an extra-time format, encourage the LSAC to modify the LSAT administration so that two scores would be offered. One would measure ability (*Mastery* score) and another would measure speed (*Intensity* score). The *Mastery* score would simply measure the percent of all questions attempted which were answered correctly, provided that a low minimum number of questions were attempted (perhaps 50% of questions offered). This *Mastery* score would give law schools a consistent measure of the logical reasoning abilities of a prospective student that was essentially independent of economic or life circumstances that impact ability to prepare for the LSAT, and independent of the number of test questions.

The *Intensity* score would be provided so that the LSAT retains its ability to reward the advantaged (and/or extremely committed) with high scores in an extremely competitive environment. This would be measured as today's score is. However, the test would be intentionally speeded to ensure significant competition among the best prepared and to differentiate between the two types of scores. To produce the speeded nature, additional questions should be asked in each section.

Depending on how the scores are weighted by individual schools, universities might be able to achieve diversity goals without adjustment of scores or compensation for them. Going to an LR-only format would make this a practical strategy that even further alleviates preparation inequities, as well as racial and ethnic testing biases.

B. Regarding helping all interested parties to understand the impact of the LSAT on individual lives – assumes that LSAT will be present in some form:

9. Help to eliminate any stigma associated with a low test-score, better arm those lacking quality preparation opportunities, and pressure law schools to de-emphasize the LSAT by announcing clearly and *regularly*, – and mandating that all law schools you accredit do the same:

“The LSAT appears to be most useful in measuring your ability to prepare for unfamiliar challenges in a short span of time, not your aptitude for legal work, your academic background or your intelligence. You will generally do poorly if your life circumstances do not permit you the time to study and the financial means to purchase preparation books, old tests, or commercial courses. Most law schools will attempt to compensate for a low LSAT score by considering other factors, but they grant few such compensatory admissions. If you believe this policy is unfair and would like to change it, please go to http://www.***.org to register your concerns, your LSAT experience, and/or your ideas for improving the system. LSAT statistical studies can be found at <http://www.lsac.org/LSACResources>. Independent evaluations of those studies can be found at http://www.***.org.”

10. Commence immediate, independent research:
 - a. Have your own statisticians review the LSAC studies to gauge their scientific merit, particularly the use of the correlation coefficient (r) as opposed to the

more informative coefficient of determination (R^2) and whether r-values as low as 0.3 – 0.5 could be considered “strong” correlations.

- b. Perform studies (and encourage LSAC to release their own studies, if any) that measure the extent to which the LSAT really measures a prospective student’s time and resource availability/commitment versus law school preparedness (ability in the absence of time), as well as whether additional testing time would better enable test-takers to demonstrate ability.
- c. Investigate why undergraduate GPA’s, according to LSAC studies, are not good predictions of law school performance. If law schools require skills not taught at the undergraduate level, then identify how the undergraduate curriculum should be supplemented or otherwise modified. If, however, law schools employ teaching methods that do not translate into good legal practice, then also propose changes in law school curriculum.

Whether you choose LSAT modification, LSAT abolition, independent research or other strategies, future generations of law students ask only that you act immediately to help improve the law school admissions process in the law schools you accredit or administer, as well as work with LSAC, where necessary. Why should another set of 150,000 prospective law school students in the next test administration year be subjected to a grueling, unnecessary and unfair process, when you have the power to chart a more accurate, more compelling and more equitable course?

3. INTRODUCTION

3.1. The LSAT and law school accreditation

LSAC describes the LSAT as follows:

The LSAT is a half-day, standardized test required for admission to all ABA-approved law schools, most Canadian law schools, and many other law schools. ... The LSAT is designed to measure skills considered essential for success in law school: the reading and comprehension of complex tests with accuracy and insight; the organization and management of information and the ability to draw reasonable inferences from it; the ability to think critically; and the analysis and evaluation of the reasoning and argument of others.¹

The *2010-2011 ABA Standards for Approval of Law Schools* states, in relevant part²:

Standard 503. ADMISSION TEST

A law school shall require each applicant for admission as a first year J.D. student to take a valid and reliable admission test to assist the school and the applicant in assessing the applicant's capability of satisfactorily completing the school's educational program. In making admissions decisions, a law school shall use the test results in a manner that is consistent with the current guidelines regarding proper use of the test results provided by the agency that developed the test. (emphasis original)

Interpretation 503-1

A law school that uses an admission test other than the Law School Admission Test sponsored by the Law School Admission Council shall establish that such other test is a valid and reliable test to assist the school in assessing an applicant's capability to satisfactorily complete the school's educational program. (emphasis original)

The November 2010 proposed modifications to Standard 503 are, simply, to withdraw it and to modify Standard 501 modestly. Those modifications are, in relevant part, indicated below.

Standard 501. ADMISSIONS

(a) A law school shall maintain sound admission policies and practices, consistent with its mission and the objectives of its educational program ~~and the resources available for implementing those objectives.~~ (emphasis original)

¹ Law School Admission Council (LSAC), *The Official LSAT SuperPrep* (Newton, PA: Law School Admission Council, 2007), 1.

² American Bar Association, *2010-2011 Standards and Rules of Procedure for Approval of Law Schools* (Chicago, IL: American Bar Association), 36, http://www.americanbar.org/content/dam/aba/migrated/legaled/standards/2010-2011_standards/2010-2011abastandards_pdf_files/chapter5.authcheckdam.pdf (accessed approximately March 1, 2011).

(b) A law school shall not admit applicants who do not appear capable of satisfactorily completing its educational program and being admitted to the bar. (emphasis original)

Interpretation 501-1

Sound admissions policies and practices may include consideration of admission test scores, undergraduate course of study and grade point average, extracurricular activities, work experience, performance in other graduate or professional programs, performance in a law school pre-admission program, relevant demonstrated skills, and obstacles overcome.

Interpretation 501-2

A law school's admission policies shall be consistent with Standards 211 and 212.

Interpretation 501-3

Among the factors to consider in assessing compliance with Standard 501(b) are the academic and admission test credentials of the law school's entering students, the academic attrition rate of the law school's students, the bar passage rate of its graduates, and the effectiveness of the law school's academic support program. Successful bar passage alone does not demonstrate compliance with Standard 501(b).

3.2. Must decouple the LSAT from the law school admissions process

3.2.1. Existing Standards

Under the current standards, the ABA has effectively sanctioned and mandated an invalid test. The LSAT is not a "valid and reliable admission test" because it fails to explain 84% - 95% of the variability in first-year law school GPA (FYA); correlations between LSAT scores and FYA have been decreasing precipitously over the past ten years; and the test is speeded. Therefore, any school that attempts to follow existing Standard 503, while simultaneously requiring students to take the LSAT, is in violation of said Standard.

Though the LSAT violates Standard 503 of the *2010-2011 ABA Standards for Approval of Law Schools*, in that it is not "valid and reliable," Interpretation 503-1 explicitly exempts the LSAT from critical review, tacitly accepting it as a valid and reliable test. Since we demonstrate that the LSAT is neither valid nor reliable, Interpretation 503-1 cannot continue to exempt the LSAT. Alternatively, Section 503 could be revised to prohibit the required use of the LSAT in any admissions decision. Clearly, in either case, if there are no valid and reliable substitutes for the LSAT, then the provisions of Standard 503 regarding an admissions test must be removed entirely, with the prohibition against the use of the LSAT and all other invalid or unreliable tests remaining.

Unfortunately, the proposed standards do not appreciate these subtleties. They simply do not go far enough in recognizing the contradiction of the ABA permitting invalid tests to continue, while allegedly requiring sound admissions policies.

3.2.2. Proposed Standards

The withdrawal of Standard 503 and the proposed marginal modifications to Standard 501 simply allow each law school to determine whether it will require its applicants to take the LSAT, but do not solve the inequities and invalid nature of the LSAT. In fact, these changes to the Standards force those seeking to improve the fairness of admissions processes, while increasing or maintaining diversity, to fight 200 separate battles. This would guarantee that many of the law schools the ABA oversees will continue to accept results that are scientifically unsound. So, merely abolishing the LSAT as an ABA mandate is not sufficient. The proposed Standard 501 must be revised to ensure that the LSAT in its current form is never permitted in any admissions process by any law school which the ABA accredits. Language is suggested in the previous section.

Regrettably, the proposed Section 501 and its Interpretation 501-1 appear to be on the fast track for passage. The ABA, therefore, will be complicit in individual law schools subjecting their applicants to an unnecessary, error-prone, discriminatory experience. However, there are steps that the ABA can take to help the LSAT or whatever standardized test is used to become a more reliable measure of ability and to reduce or eliminate the inequities associated with it. They are summarized in the previous section and explored in great detail below.

The ABA must take this responsibility seriously. If it does not, then the Comprehensive Review of the Standards initiative will have been largely a show, having accomplished little that will affect the lives of 150,000 law school applicants yearly.

We ask that you review the evidence and arguments presented in this communiqué, conduct independent validation of LSAC's studies, and compel LSAC to release relevant data and unpublished studies, if any, that paint a more accurate picture of what the LSAT measures and how test-takers actually obtain higher scores. We ask that you do similarly with other standardized tests as they become available. We ask you to publicize what you learn and modify the proposed Section 501 as necessary.

3.3. The urgency of now

I am a prospective law school student who went through an agonizing process during the 2010 – 2011 LSAT test administration year to prepare for two LSAT exams. I did poorly on the first, when I had virtually no time to prepare. I answered approximately 19 additional questions correctly³ on the second test, due solely to the fact that I had more opportunities to prepare. What I will never know is how much better I would have done if I had the time and financial resources to prepare adequately or, at least, competitively, with the more advantaged test-takers. Neither will I be able to recoup the anxiety-ridden months leading up to the tests.

My story is repeated one hundred thousand times each year. I suspect, however, that my peers are so ashamed of test scores unreflective of their respective abilities and achievements, that they simply suffer in silence. It is apparent to me, can be partially

³ I do not know the exact amount because my second test was in February 2011 and LSAC does not release raw scores from the February test administration.

demonstrated by available data, and will be fully demonstrated by supplementing the anecdotal reports with new data, that the LSAT is the problem, not the hard-working prospective students who lack the advantages of their counterparts.

It is time to end the trauma that tens of thousands of hopeful and visionary people are forced to experience every year. It is time to withdraw an admissions process that could falsely predict poor law school performance up to 95% of the time.

I have devoted an inordinate amount of time to pull together all of the relevant research I could find. When making every reasonable argument that I could in support of the abolition of the LSAT or its drastic modification, I have gone only where the evidence and my personal experience led me.

I must now return to the overwhelming demands of my life. The rest is up to you.

3.4. Organization of remaining sections - what we will show

There are seven major sections below. Typically, what we learn from our analysis of the LSAT can be applied to other standardized tests.

3.4.1. Proposal A. Modify the LSAT and its Test Administration Procedures to Reduce or Eliminate Inequities.

In this section (about 1/3 of the remaining document), we:

- Describe what can be called the *LSAT test-taker's preparation burden*, which comprises the time, money and sacrifices required for an individual to prepare fully for the LSAT.
- Present LSAC studies and other evidence that help to demonstrate the burden's existence; the percent of test-takers likely to experience an unusually high burden; inequities resulting from an unmet burden; impact of lack of preparation on LSAT scores.
- Survey the research literature and present findings that show standardized testing is invalid to the extent that it measures speed, instead of ability.
- Formulate assumptions and constraints that should be met if the LSAT were to be improved.
- Develop several improvements that meet our assumptions, while remaining within our constraints, including:
 1. Removal of the AR and RC sections;
 2. Requiring that the LSAT adopt an untimed or extra-time administration;
 3. Creation of the Mastery (power) score;
 4. Creation of the Intensity (speed) score;
 5. Exploring how law schools could utilize both scores to achieve any desired levels of selectivity and diversity, all based on merit, not gender, ethnicity or race.
- We call for additional study, but indicate that policymakers can rely upon existing research and their own experience to make changes, if that study is not forthcoming.

3.4.2. Proposal B. LSAC's studies contain errors, peculiar choices, spin that distorts the poor performance of the LSAT, and a clear refusal to investigate questions damaging to the LSAT.

In this section (about 1/3 of the remaining document), we:

- Describe errors and peculiar choices in LSAC studies – all content moved to the exhibit for readability.
- Describe the intentional, pervasive abuse of the correlation coefficient to impart a degree of validity on the LSAT which is utterly unjustified. We:
 1. Describe and define the correlation coefficient (r);
 2. Describe and define the more relevant coefficient of determination (R^2);
 3. Describe common misperceptions of the correlation coefficient that benefit LSAC and the LSAT;
 4. Describe a theory of why tiny correlation coefficients that clearly indicate non-association or minimal correlation, as measured by R^2 , seem acceptable in the behavioral sciences;
 5. Compile correlations presented by several LSAC studies over the last twenty years in tables, calculating the R^2 values that LSAC conveniently neglected to publish;
 6. Demonstrate that twenty years of LSAC studies have proven the ineptitude of the LSAT to predict FYA (84% - 95% failure rates) and that even the paltry correlations measured have been rapidly declining over the last decade;
 7. Invite policymakers to re-examine their support of the LSAT, based on renewed interpretations of correlation coefficient and their responsibility to determine what behavioral science cannot, at present, an acceptable value for r .
- Present additional LSAC spin in numerous studies where extremely poor correlation coefficients are used to reach significant conclusions that are clearly not justified.
- Formulate numerous questions – which LSAC never asked or studied – that could have determined, long ago, the relationship between demographics, preparation and LSAT performance.
- Describe a common trick by LSAT researchers in which they hide the true spread and magnitude of bad news for LSAC and the LSAT by publishing the mean, when median and percentiles would be far more informative.
- Describe the 35-point range in mean LSAT scores at different law schools and suggest that accrediting agencies should waive LSAT requirements for those schools that clearly see (low) LSAT scores as irrelevant to qualifications or matriculation.

3.4.3. Proposal C. LSAT's predictability rates of FYA justify rescinding requirements that students take the LSAT and recommend a 100% compensatory model.

In this brief section, we describe the latest national LSAC study which calculates the national mean correlation, $r = 0.33$ (and $R^2 = 0.11$). We argue that an exam that is essentially wrong 89% of the time should not be required of any applicants. We cite one study's description of compensatory admission models as ones in which applicants "are admitted because the school has some other [non-LSAT] evidence of their ability to do well in law school." We ask why all admissions decisions could not be made similarly.

3.4.4. Proposal D. Racial and ethnic LSAT disparities must be eliminated within the next twenty years or diversity efforts may be ruled unconstitutional.

In this section, we draw from *Bakke* and *Grutter* to recapitulate the Supreme Court's determination that diversity in law school is a compelling state interest, but that the 25-year time-bomb set by Justice O'Connor in *Grutter* will go off within the next twenty years. We capture LSAT scores from 1993 and from 2008 to demonstrate that there has been very little narrowing of the gap between underrepresented people of color and either Whites or Asians. We postulate that another twenty years will not make a difference either. We point out that there must be dramatic change in the LSAT *immediately*, so that it does not differentiate racially and ethnically, or current diversity efforts will be deemed unconstitutional, without recourse. We note that all of the recommendations made in this communiqué were made to lessen the inequities associated with the LSAT test-taker's burden or to address failures in the LSAT itself. All of those recommendations could be adopted to benefit underrepresented people of color and, where needed, females, just as well. We list the key recommendations.

3.4.5. Proposal E. Ways to reduce LSAT inequities, other than modification or abolition.

In this section, we describe ways that, in the absence of changes to the LSAT, we could lessen some of the inequities associated with the LSAT. These include independent studies by interested parties; public disclosures about what the LSAT is and is not; high-quality, free LSAT test preparation classes sponsored by universities. Interested parties are encouraged to setup websites to invite feedback from test-takers to better assess the utility of the LSAT.

3.4.6. Conclusion

In this section, we summarize the purpose, findings and recommendations of this communiqué, then issue a call to action.

3.4.7. Exhibit A

This section contains a catalog of errors and peculiar choices made by LSAT researchers, as well as a relevant statistics calculations primer in the footnotes. It was removed from the main body of the document for readability purposes. What we uncovered includes the following:

- In the *Analysis of Differential Prediction of Law School Performance*, series:
 1. There are attempts to compare the mean of various statistical measures using non-standard statistical choices and there is no reliance upon the *t*-test or other proven techniques;
 2. Averages of correlation coefficients were taken, producing an unknown and undefined statistic;
 3. The correlation coefficients for each subgroup were calculated separately by law school, then averaged for all schools in an effort to capture differences in the national performance of subgroups. Correlation at the university level was unnecessary, since grades had been normalized. This

distorted the actual correlations per subgroup and created an undefined statistic as indicated above;

4. Inexplicably, the “weighted average of the mean residuals between predicted and actual FYA” was calculated, instead of the mean of the squares of the residuals or the mean of the absolute values of the residuals. This allowed some of the residuals (negative vs. positive values) to cancel out, making the errors appear much smaller than they really were;
 5. Numerous mathematical operations were performed on approximations of non-linear statistics, producing undefined statistics; subpopulation statistics were calculated using regression equations not solved specifically for that particular subgroup; etc. Conclusions drawn from these statistics are of dubious value.
- In the *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*:
 1. Researchers formed the mean of the correlation coefficients which were calculated for each law school, without regard to the different variances, distributions or any other different characteristics of the schools;
 2. Researchers defined “selected group” and “unselected group” such that the former was a part of the latter, though their names convey mutual exclusivity. This conflation made its way into the calculations;
 3. The selection effect calculations appeared to be in error.
 - In *LSAT Item-Type Validity Study (TR 98-01)*:
 1. The raw LSAT score was used, when the scaled score was the standard; we surmised that the researchers chose an unusual data set, for no methodological reason, because it simply generated “better” results for them;
 2. The researchers used an inferior standard for repeat test-takers, even though previous LSAC studies informed them of this fact.
 - Similar errors and peculiar choices manifested themselves in other studies.
 - We noted that qualified academics should review these studies and our criticisms to determine whether and how much cause for alarm there is.

4. PROPOSAL A. MODIFY THE LSAT AND ITS TEST ADMINISTRATION PROCEDURES TO REDUCE OR ELIMINATE INEQUITIES.

4.1. Today's LSAT

The LSAT currently consists of four scored sections, each timed for 35 minutes, and a writing sample, also timed for 35 minutes.⁴ They are:

Reading Comprehension (RC)	26 – 28 items
Logical Reasoning (LR)	24 – 26 items
Logical Reasoning (LR)	24 – 26 items
Analytical Reasoning (RC)	22 – 24 items

Candidates may request test accommodations that differ from the above, with respect to being allowed additional time to complete the exam or for any other adjustments, provided the reason for the requested accommodation is a disability.⁵ Decisions on accommodations are made on a case-by-case basis, after appropriate documentation is supplied to LSAC.⁶

4.2. Why test-taker preparation is necessary for the LSAT

As noted above, according to LSAC, the LSAT is “designed to measure skills considered essential for success in law school.” As such, it attempts to measure skills acquired over time. Florida Atlantic University states:

The LSAT is not an IQ test. It does not measure intelligence the way IQ tests are designed to measure innate ability. A person who is very smart can receive a low LSAT score. A person who is very creative can receive a low LSAT score. Furthermore, because an IQ test presumably measures natural / innate rather than learned/acquired ability, a person cannot study or prepare for an intelligence test the way that a person can study, prepare for, and improve performance on the LSAT.⁷

Unfortunately, the LSAC's advice is so understated that it simply cannot be seen as truthful or helpful:

Most law school applicants familiarize themselves with test directions and question types, practice on sample tests, and study the information available on test-taking techniques and strategies. Although it is difficult to say when

⁴ Susan P. Dalessandro, Lisa A. Stilwell, Jennifer A. Lawlor and Lynda M. Reese, *LSAT Performance with Regional, Gender, and Racial/Ethnic Breakdowns: 2003–2004 Through 2009–2010 Testing Years*, LSAT Technical Report 10-03 (Newton, PA: Law School Admission Council, October 2010), 4, <http://www.lsac.org/LsacResources/Research/TR/TR-10-03.pdf> (accessed approximately March 1, 2011).

⁵ LSAC, “The LSAT: Accommodated Testing,” <http://www.lsac.org/JD/LSAT/accommodated-testing.asp> (accessed approximately March 1, 2011).

⁶ LSAC, “The LSAT: Accommodated Testing.”

⁷ Florida Atlantic University, “The LSAT,” <http://www.fau.edu/prelaw/lSAT.php> (accessed approximately March 1, 2011).

examinees are sufficiently prepared, very few people achieve their full potential without some preparation.⁸

Negligently understated though it is, by its own admission, the LSAT requires preparation on the part of its candidates.⁹

4.3. The LSAT test-taker's preparation burden

According to LSAC surveys, less than 3% of LSAT test takers use no test preparation method whatsoever.¹⁰ However, approximately¹¹:

- 97% of all test-takers utilize at least one test preparation method.
- 1/2 of all test-takers use *two or more* test preparation methods.
- 1/3 of all test-takers utilize *three or more* test preparation methods.
- 1/3 of all test-takers use a commercial test-preparation school and 40% use a book *not* published by the LSAC.¹²

To continue examining preparation required, it is helpful to discuss what might be called the *LSAT test-taker's preparation burden*. We can define that *burden* as:

The *time* necessary to prepare for the LSAT, the amount of *money* required to purchase materials and register for a course, and the personal/professional *sacrifices* required to complete that study.

The *LSAT test-taker's preparation burden* is the source of the inequities that the LSAT imposes and which is the fundamental concern of this communiqué.

4.3.1. Required: Time

While the LSAC studies do not indicate the amount of time typically required for preparation, universities and commercial test preparation schools do. The most common calendar duration for test preparation is three months. Howard University, for instance, tells its prospective students this on its admissions page¹³:

We recommend that you take the LSAT early, allowing at least 3 months to prepare. Remember, the LSAT is not a test of your intelligence but a test of your preparedness.

⁸ LSAC, "The LSAT: Preparing for the LSAT," <http://www.lsac.org/JD/LSAT/preparing-for-LSAT.asp> (accessed approximately March 1, 2011).

⁹ There are anecdotal accounts of individuals who do not prepare, but still do well. That type of intellectual or test-taking genius is clearly rare and outside the scope of this communiqué.

¹⁰ Josiah Evans, Andrea E. Thornton and Lynda M. Reese, *Summary of Self-Reported Methods of Test Preparation by LSAT Takers for Testing Years 2005–2006 through 2007–2008, LSAT Technical Report 08-04* (Newton, PA: Law School Admission Counsel, October 2008), 11, Table 10, <http://www.lsac.org/LsacResources/Research/TR/TR-08-04.pdf> (accessed approximately March 1, 2011).

¹¹ Evans, *Summary 08-04*, 11.

¹² Evans, *Summary 08-04*, 9, Table 7. Note that this LSAC data was based on voluntary responses, so there is the possibility that the use of commercial test-preparation schools is significantly under-reported due to a potential stigma of having to "cheat" or obtain an unfair advantage to get into law school.

¹³ Howard University, "FAQS (Frequently Asked Questions)," <http://www.law.howard.edu/518> (accessed approximately March 1, 2011).

The commercial test-preparation schools provide similar advice and succinct direction is provided by the popular web site Top-Law-Schools.com at <http://www.top-law-schools.com/score-well-on-lsat.html>:

In order to get the true benefit of the class, you have to do all of the assigned homework. While there is usually a lot of homework, as discussed above, the best way to improve your score is through practice. ... Some prefer to do only practice tests and nothing else to prepare and do upwards of 30 practice tests. ... For most people, it takes at least **two or three months of studying** to fully prepare for and feel confident about the LSAT. Whether you study for one month or six, the key is to make sure that you are studying consistently. (emphasis added)

The amount of time actually spent in preparation is more telling. PowerScore provides a “full-length course” with more than 125 lecture hours, including 13 four-hour lecture sessions that meet over an eight week period.¹⁴ Additional after-hours study is required. One LSAT tutor has published study schedules including a four-month schedule of 15 hours per week and a three-month schedule of 20 hours per week.¹⁵ Other tutors have similar advice.¹⁶

Everyone has a significant responsibility to prepare for the LSAT. However, this becomes unduly and inequitably burdensome when a test-taker simply cannot obtain the time required due to life circumstances (not sloth).

4.3.2. Required: Money

The cost for a commercial course is extraordinary. PowerScore charges \$1,295 for its full-length course. Kaplan’s classroom instruction ranges from \$1,399 (33 classroom instructional hours and 16 testing hours spread across 14 sessions)¹⁷ to \$7,499 (six weeks of an intensive summer “camp”)¹⁸.

¹⁴ PowerScore Test Preparation, “The PowerScore Full-Length LSAT Course - \$1,295,” http://www.powerscore.com/lSAT/full/content_full.cfm (accessed approximately March 1, 2011) and PowerScore Test Preparation, “Which LSAT Course Should I Take?,” http://www.powerscore.com/lSAT/content_whichcourse.cfm (accessed approximately March 1, 2011).

¹⁵ Steve Schwartz, “LSAT Prep While Working or in College,” August 20, 2010, <http://lsatblog.blogspot.com/2010/08/lSAT-prep-while-working-college.html> (accessed approximately March 1, 2011).

¹⁶ See <http://www.lsatprep.org> (accessed approximately March 1, 2011), <http://lawschoolinteractive.com/category/lSAT-prep> (accessed approximately March 1, 2011), Nova Press – Test Prep Center, “LSAT Prep to fit your lifestyle and budget,” <http://novapress.net/lSAT/lSAT-prep-materials> (accessed approximately March 1, 2011), and the other LSAT sites listed in this document.

¹⁷ Kaplan Test Prep, “LSAT Advantage – On Site,” <http://www.kaptest.com/LSAT/LSAT-Prep/Classroom-On-Site/lSAT-advantage-on-site.html> (accessed approximately March 1, 2011).

¹⁸ Kaplan Test Prep, “LSAT Classroom On Site options,” <http://www.kaptest.com/LSAT/Home/classroom-on-site-options.html> (accessed approximately March 1, 2011).

Books not published by LSAC, one category of the most popular study materials, cost approximately \$50 each.¹⁹

Everyone has a responsibility to purchase materials for the LSAT and/or pay for course registration, depending on preparation choice. However, this becomes unduly and inequitably burdensome when a test-taker simply cannot obtain the funds required for such purchases (for reasons other than frivolous or unwise spending choices).

4.3.3. Required: Personal/Professional sacrifices

The measure of any time or financial preparation burden can sometimes be determined, trivially, by whether a prospective student actually has the time in his day or the funds to obtain the materials/register for preparation classes, respectively. However, burdens could still exist when there are extraordinary demands on one's time or when funds that are physically present might seem better spent elsewhere. This is likely what occurs in the great majority of cases. Hence, the extent or magnitude of a preparation burden can often be measured only when one looks at the types of sacrifices or tough choices that must be made if one were to devote the requisite time and/or money to prepare for the LSAT. Below, we describe some of these.

It is difficult to gauge the types of meaningful sacrifices required of typical young, single, childless, college students or recent graduates – still receiving financial support from parents – to meet the above time and monetary commitments to prepare for the LSAT. It is clearly an individual matter. Perhaps time taken from looking for employment searches or money spent on LSAT preparation in lieu of personal expenses is relevant in this regard.

However, as we successively withdraw each of the above attributes (age, marital status, childlessness, support from parents) and add typical others, we can see that the personal and professional sacrifices required for LSAT preparation can easily become overwhelming or insurmountable. Even a well-organized LSAT candidate, who planned to have all of his affairs in order by the time law school commences, may yet face these challenges during the months before test time.

The sources of sacrifice, then, include:

- Having personal or family responsibilities that require all available income, just to survive – This almost certainly guarantees that no monetary resources will be devoted to preparation for the LSAT;
- Need to care for children or the elderly – This reduces the time and money available to prepare for the LSAT;
- Pressing or emergency spousal or other family needs – Impacts time and money that would otherwise be available for LSAT prep;
- Career responsibilities – Projects, deadlines, personnel issues, sales calls, client issues can each or all interfere with prep time;

¹⁹ See, for instance, the three listed at PowerScore Test Preparation, “Self-Directed LSAT Study,” http://www.powerscore.com/lsat/content_index.cfm (accessed approximately March 1, 2011).

- Health issues²⁰ – There could be chronic health issues that do not rise to the level of a typical disability, but which are financially taxing or which reduce preparation time. These include complications from arthritis, diabetes, hypertension, back pain;
- Self-employment – Impacts both time (operating a business typically leaves no time for anything else, as well, there are no sick or vacation days available) and money (if the business is struggling, there is no money and what funds are present pay employees and creditors first, not the self-employed individual or his LSAT fees).

4.4. Inequities resulting from the *LSAT test-taker's preparation burden*

Any LSAT test-taker who has any of the above attributes or faces any of the above challenges would have to choose between preparing for the LSAT and managing the affairs of his/her life. The difficult choices required could include the following:

- Do I pay for a babysitter or pay for an LSAT prep book? If I buy the book, can I study at home with a young child demanding attention?
- How do I divide my time between caring for my elderly parent, juggling work assignments and preparing for the LSAT?
- Does my role as a self-employed businessperson *ever* allow me a single block of time to take a practice exam (this is recommended universally as a critical aspect of preparation)?
- If I pay for a prep course, how will I pay my mortgage, rent, car payment, or eat?

Clearly, these challenges and choices become magnified with age and with diminished income. Just as likely is that the pressing life issues will take priority and LSAT preparation will simply not occur.

And, yet, there is only one LSAT score and those with the above disadvantages are competing against those with none of these disadvantages. Inequities in test preparation should easily translate to less test prep time and, axiomatically, a lower test score.

4.4.1. Pervasiveness and intensity of inequities from *LSAT test-taker's preparation burden*

The question of the magnitude of the problem of test inequities is critical because it could determine the magnitude of the response from accreditation agencies, LSAC and even oversight bodies, such as the US Department of Education.

We can begin by asking whether there is any evidence that *some* prospective law school students do not have the time and/or financial resources to prepare for the exam as well as others. Of course, we already know that the answer to this question is emphatically yes, because the LSAC fee waiver process exists and we know that poverty exists everywhere.

²⁰ It might be appropriate to include all physical and mental disabilities as part of the challenges that have to be overcome in LSAT prep. However, disability accommodations are already the right of any disabled applicant. That the accommodations are reported to law school and score percentiles are not supplied by LSAC is another matter. Fortunately, some of the suggested modifications can be applied to helping those with disabilities. But, because LSAC has already taken them out of the scoring and hardship equation, to a large degree, factoring them back into this broader discussion would not be helpful to either the disabled (trivializing their concerns) or non-disabled (whose problems and challenges might pale in comparison).

It can be safely assumed that those receiving fee waivers cannot afford the commercial test preparation schools or even the non-LSAC books popular among the survey respondents.

Secondly, it seems nearly certain that several categories of LSAT candidates will have characteristics highly correlated with those requiring the sacrifices identified above: those below certain income ranges; those who are from underrepresented groups (greater incidences of poverty, teenage pregnancy, single parenting, etc.); older test-takers (more likely to have families); females (perhaps likely to have childcare or eldercare issues), etc. All of these categories of test-takers exist in significant numbers, so the effect of inequities must also be significant.

There is some data that allows us to make a rough attempt at quantifying this. According to LSAC's analysis of 2010 data, the number of underrepresented people of color and females who took the LSAT in the 2008 – 2010 test administration years is as follows.²¹

²¹ For race/ethnicity and gender data (2009 – 10 test year), see Dalessandro, *et. al.*, *LSAT Performance*, 9, 16 and 20. For age data (2008 – 2009 test year), see Kimberly Dustman and Phil Handwerk, *Analysis of Law School Applicants by Age Group: ABA Applicants 2005-2009* (Newton, PA: Law School Admission Council, October 2010), 3, <http://www.lsac.org/LSACResources/Data/PDFs/Analysis-Applicants-by-Age-Group.pdf> (accessed approximately March 1, 2011).

Table 1: Populations Likely Impacted by Excessive LSAT Test-Taker’s Preparation Burden, 2008 – 10

from *LSAT Performance with Regional, Gender, and Racial/Ethnic Breakdowns: 2003–2004 Through 2009–2010 Testing Years (TR 10-03)* and *Analysis of Law School Applicants by Age Group: ABA Applicants 2005-2009*

Subgroup	Number taking the LSAT
Age:*	
30 – 39	11,635
>= 40	4,515
Subtotal	16,150
TOTAL test-takers	89,251
% of Total test-takers	18%
Gender:**	
Female	60,651
TOTAL test-takers	125,331
% of Total test-takers	48%
Race/Ethnicity:**	
American Indian/Alaskan Native	634
Black/African American	14,585
Canadian Aboriginal	50
Native Hawaiian/Other Pacific Islander	340
Hispanic/Latino	9,264
Puerto Rican	2,347
Subtotal	27,220
TOTAL test-takers	125,300
% of Total test-takers	22%

*Age data is from 2008 – 9. Percent & total test-taker numbers based on those who answered voluntarily.

**Race/ethnicity and gender data is from 2009 – 2010. Percent & Total test-taker numbers based on those who answered voluntarily.

This data indicates that the numbers of people likely to be disproportionately and negatively impacted by their *LSAT test-taker’s preparation burden* are, by age, 18%, by race/ethnicity, 22%, by gender, 48%.²² Note that the numbers of Whites and Asians similarly impacted (such as those who are financially poor), cannot be inferred from LSAC’s data. The above numbers are so significant, however, that they suggest an inordinate *LSAT test-taker’s preparation burden* might be the rule, not the exception.

²² These percent calculations are based on the responses of those who *voluntarily* responded.

LSAC is in a strong position to provide data on this topic by releasing information on what percentage of its registrants receive fee waivers. Additional direct study would also be extremely helpful. However, it is clear that the effect of *LSAT test-taker's preparation burden* may be pervasive and must be examined more carefully.

4.5. The *LSAT test-taker's preparation burden* and LSAT scores

As indicated above, we would expect that a greater *LSAT test-taker's preparation burden* should yield lower LSAT scores. Unfortunately, while we can intrinsically see this relationship, there seems to be no data directly measuring it. However, there is a degree of empirical and other evidence.

4.5.1. LSAC Research: LSAT scores decline with age – implications for preparedness

There is one study that strongly suggests that challenging life circumstances deny certain test-takers sufficient time to prepare and, therefore, they achieve a lower test score than might otherwise be the case. The study is *Analysis of Law School Applicants by Age Group: ABA Applicants 2005-2009*.²³ The researchers show that mean LSAT scores decline precipitously with age, from a mean of 155 for the 25 and under subgroup to a mean of 144, for the 40+ subgroup.²⁴ The physiology of aging could be responsible for the decline, though such a significant drop in performance seems extremely unlikely, based on current research.²⁵

However, there is a more reasonable interpretation. As one ages, she is more likely to have family responsibilities and face financial pressures. Therefore, the older prospective student is likely to lack both the preparation time and money (for study materials and commercial test preparation classes) that the younger test-taker has. Hence, the older student is likely to be less prepared for the LSAT and would perform more poorly. The

²³ Dustman and Handwerk, *Analysis*, 8.

²⁴ There could be other factors that partially explain the data, such as the greater number of African-American test-takers among older subgroups. (African-Americans consistently have a lower performance than most other groups for reasons that the LSAC has never explained.)

²⁵ See Florida State University, "Adult Learning," <http://www.fsu.edu/~adult-ed/jenny/learning.html> (accessed approximately March 1, 2011), Timothy A. Salthouse, "New Study Reconciles Conflicting Data on Mental Aging," September 13, 2010, <http://www.apa.org/news/press/releases/2010/09/mental-aging.aspx> (accessed approximately March 1, 2011) and Timothy A. Salthouse, "Influence of Age on Practice Effects in Longitudinal Neurocognitive Change," *Neuropsychology* 24, No. 5 (2010): 563 – 572, <http://www.apa.org/pubs/journals/releases/neu-24-5-563.pdf> (accessed approximately March 1, 2011). See also Sharan B. Merriam, Rosemary S. Caffarella and Lisa M. Baumgartner, *Learning in Adulthood: A Comprehensive Guide / Edition 3* (Somerset, NJ: Wiley, John & Sons, Incorporated, 2006), Chapter 8, <http://search.barnesandnoble.com/Learning-in-Adulthood/Sharan-B-Merriam/e/9780787975883> (accessed approximately March 1, 2011):

Does intelligence decline with age? Responses to this question are mixed and often controversial. The classic school of thought contends that intelligence enters a process of irreversible decline in the adult years, although the hypothesized onset of that decline has been extended from the early twenties to at least the age of fifty or sixty. Others say that intelligence is relatively stable through the adult years, with substantial intellectual changes occurring only very late in life, and then primarily "in abilities that were less central to the individual's life experience and thus perhaps less practiced" (Schaie, 1996, p. 2). In essence, we have enough brain capacity to do almost anything we choose, until serious illness sets in. Still others argue that intelligence declines in some respects, remains stable in others, and may even increase in some functions

data LSAC already possesses, and certainly additional data, could routinely confirm (or refute) this.

4.5.2. LSAC Research: LSAT Repeater Data

There is a file that can be difficult to locate on the LSAC web site, even if you know its name.²⁶ It is called “RepeaterData.pdf” and is available on the LSAC site at <http://www.lsac.org/LSACResources/Data/PDFs/repeaterdata.pdf>. This file tracks test scores of those who repeat the LSAT. For each possible previous test score, it lists the number of test-takers whose new scores are within successive ranges of approximately 10, from 120 – 129, 130 – 139, through 170-180. The current file shows data for those who repeated the LSAT in the 2009 – 2010 test administration year, after having taken it at some point in the past.²⁷ Unfortunately, because the range of new test scores is fixed at approximately 10 scaled points, granular analysis is not possible. However, by closely examining the data, we can provide the following two tables.

Table 2: Improvements in LSAT Scores among Repeat Test-Takers
from 2009 – 10 LSAT test administrations

<http://www.lsac.org/LSACResources/Data/PDFs/repeaterdata.pdf>

Old LSAT Score	# Test-Takers Achieving Better Score: 4 – 13 Points*	Percent of Test-Takers Achieving Better Score: 4 – 13 Points*	# Test-Takers Achieving Better Score: 8 – 17 Points*	Percent of Test-Takers Achieving Better Score: 8 – 17 Points*
122	*	*	52	18%
126	265	36%	*	*
132	*	*	188	12%
136	748	38%	*	*
142	*	*	184	11%
146	437	39%	*	*
152	*	*	92	14%
156	142	42%	*	*
162	*	*	52	18%
166	265	36%	*	*

*Range provided by LSAC was 10 scaled points: 130-40, 140-50, 150-60, 160-70, 170-80. No other data was made available. Note that the two test score improvement ranges (4 – 13 and 8 – 17) do not actually overlap, because they refer to different portions of the data.

We can translate this data into the *number of additional correct answers* that each of these new scores represented, per LSAC’s score conversion tables.²⁸

²⁶ The ease with which you can find this file seems to depend on which search box you use. It can be found using the link <http://www.lsac.org/LSACResources/Data/PDFs/repeaterdata.pdf> or from the “Data” link on the “LSAC Resources” tab.

²⁷ LSAC, “LSAT Repeater Data,” <http://www.lsac.org/LSACResources/Data/PDFs/repeaterdata.pdf> (accessed approximately March 1, 2011).

²⁸ This data is created by LSAC and distributed to students with their results during certain test administrations. Many sources are available online. The above data points are taken from Alpha Score

Table 3: Number of Additional Correct Answers for Each Improved Score Range
2009 – 10 LSAT test administration

from <http://www.lsac.org/LSACResources/Data/PDFs/repeaterdata.pdf> and
<http://www.alpha-score.com/resources/lsat-score-conversion>.

Old LSAT Score	# Additional Correct Answers for Better Score: 4 – 13 Points*	Percent of Test-Takers Achieving Better Score: 4 – 13 Points*	# Additional Correct Answers for Better Score: 8 – 17 Points*	Percent of Test-Takers Achieving Better Score: 8 – 17 Points*
122	*	*	8 - 19	18%
126	4 - 15	36%	*	*
132	*	*	10 - 25	12%
136	6 - 21	38%	*	*
142	*	*	13 - 29	11%
146	7 - 23	39%	*	*
152	*	*	14 - 29	14%
156	7 - 22	42%	*	*
162	*	*	13 - 24	18%
166	6 - 17	36%	*	*

*Range provided by LSAC was 10 scaled points: 130-40, 140-50, 150-60, 160-70, 170-80. No other data was made available. Note that the two test score improvement ranges (4 – 13 and 8 – 17) do not actually overlap, because they refer to different portions of the data.

In the samples above, chosen from non-overlapping sections across the entire score range, but for no particular attribute, 39% of repeat test-takers improved their most recent scores by 4 – 13 scaled points (4 – 23 additional correct answers) and 13% improved their most recent scores by 8 – 17 scaled points (8 – 29 additional correct answers).²⁹

These performance gains from an additional LSAT are dramatically different from the mean figures typically provided by LSAC.³⁰ We, therefore, ignore LSAC’s mean calculations, which could be distorted by outliers and which do not convey spread. Instead, we note that for 2,171 out of 8,424 repeat test-takers in our sample above to achieve significant improvements, reasons unrelated to the relenting of mass illness or a sudden rise of good fortune must be involved.³¹ We may infer that additional preparation

Online LSAT Test Preparation, “LSAT Score Conversion,” <http://www.alpha-score.com/resources/lsat-score-conversion> (accessed approximately March 1, 2011).

²⁹ Regrettably, LSAC provides no further granularity and little more can be inferred.

³⁰ LSAC measured the mean scaled score gained from a second LSAT exam to be 2.8 points and 1.9 points for a third LSAT exam. See Andrea E. Thornton, Laura A. Marcus, Arlene Amodeo and Lynda M. Reese, *The Performance of Repeat Test Takers On the Law School Admission Test: 2000–2001 Through 2006–2007 Testing Years*, *LSAT Technical Report 08-01* (Newton, PA: Law School Admission Council, March 2008), 1, <http://www.lsac.org/LSACResources/Research/TR/TR-08-01.pdf> (accessed approximately March 1, 2011).

³¹ The raw data is from LSAC, “LSAT Repeater Data.” Also from “LSAT Repeater Data,” LSAC advances several reasons for why a prospective student would want to re-take the LSAT, none of which is additional preparation time:

was the most likely factor that changed. That additional training could transform the prospects of thousands of law school hopefuls clearly challenges any LSAC notion that the LSAT consistently measures reasoning and reading comprehension ability.

The important point regarding these two studies is that there is empirical evidence that appears to support our principal argument that performance on the LSAT is related strongly to the *LSAT test-taker's preparation burden*, not some consistent measure of ability or drive to perform well in law school.³² Of course, further study is necessary.

4.5.3. Anecdotal evidence

4.5.3.1. *Repeat test-takers*

The anecdotal evidence that would be most helpful in showing a link between LSAT test-taker's preparation burden and test scores must control for all of the moving variables: the individual, the (lack of) preparation, the scores, evidence that scores were tied to level of preparation, etc. The most obvious examples would be from repeat test-takers who can demonstrate causal relationships between levels of preparation, issues affecting those levels of preparation, and their scores. Interviews with them would be cornerstones of any scientific research into how the LSAT test-taker's preparation burden affects scores.

LSAT and law school blogs contain first- and second-hand accounts that meet these criteria. Frequently, however, the comments are anonymous and we cannot discern when they are real and when they are for marketing purposes. Interviews with LSAT trainers, commercial test preparation administrators would be similarly helpful and similarly tainted.

4.5.3.2. *Personal account*

My personal story provides a fascinating glimpse into the relationship between *LSAT test-taker's preparation burden* and LSAT scores. In 2005, I received the lowest standardized test score of any kind when I took the LSAT cold. I had no meaningful knowledge of the question types and I do not believe that I had attempted a single LSAT sample question. I thought I would do well because I bought the LSAC myth that the exam merely tested your ability to do well in law school. I scored in the 90+ percentile on a GRE exam taken under identical conditions several years earlier.

I took the December 2010 exam after attending a discount weekend course three weeks before the exam. I obtained the funds by not paying other bills. I had attempted to do self-study before and after that course. But, family crises prevented that. I had completed a few drills and portions of two sample LSAT exams before test day. I had not paid much attention to test-taking skills. I received precisely the same test score as I did in 2005.

Individuals need not take the LSAT more than once unless they believe some circumstance, such as illness or anxiety, prevented them from performing as well as they might have expected.

³² The test could also be measuring other factors faced by older populations, such as less attentiveness during their allotted preparation time due to family demands and distractions, deteriorating health (fading eyesight, hearing, etc.), nagging concerns about whether law school is even a viable option for them. We do not know which, if any of these factors impact the older student, of course, because the LSAC has not studied them and accrediting bodies have not performed their own, independent research.

I actually knew much more about the exam that second time. But, I did not come close to finishing any section, just as in 2005. And, I neglected to employ virtually any of the test-taking skills I had learned. Based on what I have since learned, it is obvious that the test was significantly speeded for me and that it was invalid as a measure of my ability. In fact, I answered correctly about 85% of the time (excluding final minute guessing). I just never saw numerous questions.

Once the scores came back, I decided to take the LSAT again in February 2011. This time, and with only a month to prepare, I still could not clear a block of time that would allow me to get through a full prior/practice exam. But, I completed a few dozen drills and, by the end of my preparation was more confident than before. I only feared one thing – time. For that third exam, I received a much better score, having answered approximately 19 more questions correctly than I did two months earlier. I still ran out of time, but I had read and answered many more questions.

Though my score was suddenly respectable, I was clearly no better prepared for law school. I was just a little faster and employed good test-taking skills. *From my own experience, I concluded that the LSAT was only measuring my speed, not my ability and that the test could no more predict my performance in any university than a roll of the dice.*

4.6. Modifying the LSAT

4.6.1. Evidence, assumptions and action

As demonstrated above, there is evidence (the age study, test repeater data, anecdotal and the experience of individual test-takers) that the *LSAT test-taker's preparation burden* exists, inequities associated with it exist, those inequities directly impact LSAT test scores, and the resulting lower test scores could be impacting more than half the test-takers every test administration.

It is also clear that additional study must be conducted and that LSAC already has some of the requisite data, particularly fee waiver information.

I fully acknowledge that scientific research would have to be conducted before any recommendation indicated in this communiqué will be adopted in whole or in part, or before any modifications to the LSAT whatsoever will be made. And, that seems prudent. However, because the stakes are so high for every test-taker, complaints of LSAT inequities and validity (see much more below) are pervasive and the evidence presented above at least raises the presumption that low test scores for large subgroups of test-takers result from LSAT inequities, we are obliged to begin exploring modifications immediately. They can be fully evaluated once the research is complete. And, if no further research occurs, conferences of knowledgeable, interested parties could yield a consensus, based on their collective experience and any existing research, about whether and how to modify the LSAT.

First, we make explicit the assumptions and constraints necessary to continue this inquiry. Then, we review the relevant standardized testing research. Finally, we propose modifications to the LSAT that are consistent with current thinking on educational testing, follow from our assumptions, and are well within any appropriate constraints.

4.6.2. Assumptions and constraints

First, we accept as true the evidence and analysis presented above, namely:

- The *LSAT test-taker's preparation burden* exists;
- Inequities associated with it exist;
- Those inequities directly impact LSAT test scores; and
- Resulting lower test scores impact an unacceptably large number of test-takers, with each and every test administration.

Second, we assume that the ABA, all other accrediting bodies, law schools, US Department of Education, Congress and all law school applicants expect and demand a *fair test* that does not penalize the poor, those with difficult life challenges, older applicants, underrepresented people of color, women, the disabled or anyone else.

Third, we presume that the test should not penalize those who prepared extensively for it and that law schools may desire to retain at least one measure that allows them to differentiate among levels of preparedness or high levels of attainment.

These assumptions, essentially, state that there must be something for everyone in the new LSAT. The new LSAT must minimize test preparation inequities, while consistently measuring ability, and retaining the intrinsic quality of differentiating among the best prepared.

4.6.3. Scientific research

It is clear that the above assumptions would shape the LSAT in dramatic and predictable ways. Critical to understanding the vector of these changes is a review of the relevant research. Particularly, we need to discern how measuring an applicant's ability can be reliably accomplished when test-takers are differentiated in whether they understand a concept at test time, which inferences they can draw from that knowledge, how rapidly they can access and apply that knowledge -- all under the extreme pressure of time.

4.6.3.1. *Ability is independent of speed, extra time tends to increase scores*

The pioneering work of B. Baxter (in 1931), C.T. Myers (in 1952), F.M. Lord (in 1956) and W.G. Mollenkopf (in 1960) are now well within the mainstream of thought in educational testing research.³³ Through tests conducted on adults (Baxter, Myers and

³³ The age of these studies made them inaccessible to me through the limited online databases I was able to use. Instead, the above analysis is based on interpretations in more recent works. The first is Gerald Tindal and Lynn Fuchs, "A Summary of Research on Test Changes: An Empirical Basis for Defining Accommodations," 17 – 18, <http://www.specialed.us/discoveridea/topdocs/msrrc/Tindal&Fuchs.PDF> (accessed approximately March 1, 2011). The second is Thomas E. Brooks, Betsy J. Case and Michael J. Young, *Timed Versus Untimed Testing Conditions and Student Performance* (San Antonio, TX: Pearson Education, Inc., 2003, 2004), 3 – 6, <http://www.pearsonassessments.com> (accessed approximately March 1, 2011). Another is Betsy J. Case, *It's About Time: Stanford Achievement Test Series, Tenth Edition*

Lord) and on high school students (Mollenkopf), they demonstrated that scores on timed tests are a result of two factors, speed (rate-of-answer) and ability, but that the validity of the ability measure depends on whether sufficient time has been allowed for a demonstration of those abilities to take place.³⁴ Baxter showed that ability, as measured by power (accuracy or percent correct)³⁵ was most correlated with classroom grades (better students demonstrated greater power).³⁶ Meyers showed that ability and speed are actually orthogonal (independent).³⁷ Mellenkopf demonstrated that providing additional time to a fixed-time test benefited arithmetic scores significantly more than verbal scores.³⁸

Since 1931 there has been limited high-quality research on timed vs. untimed/extra-time standardized testing, but the results have generally been as summarized above.³⁹ Two recent large-scale studies are insightful, however.

In a 2003 Pearson Education, Inc. study of 360,000 students in grades K – 12,⁴⁰ where achievement in mathematics, the sciences, vocabulary, language, etc., was measured,⁴¹ an untimed test produced higher test results than timed tests through grade 6, but that reversed above grade 6.⁴² However, Pearson did not attribute this to a failure of untimed testing. It argued that Pearson had simply learned to allow the proper amount of time on its fixed-time exams, so an untimed test offered no greater opportunity to demonstrate ability.⁴³

(Stanford 10) (San Antonio, TX: Pearson Education, Inc., April 2003 (Revision 3, January 2004)), 3, <http://www.pearsonassessments.com> (accessed approximately March 1, 2011).

³⁴ From Tindal and Fuchs, *Summary of Research*, Baxter measured college students in R.O.T.C. (p. 21); Myers measured 600 midshipmen at the US Naval Academy (p. 25); Lord measured 649 students entering the U.S. Naval Academy (p. 24), Mellenkopf measured two groups of high school students (. 24). Discussion is on pp. 17 – 18.

³⁵ For definition of power and speededness, see Susan Ellerin Rindler, “Pitfalls in Assessing Test Speededness”, *Journal of Educational Measurement* 16, No. 4 (Winter, 1979), 261.

³⁶ Tindal and Fuchs, *Summary of Research*, 17, 21.

³⁷ Tindal and Fuchs, *Summary of Research*, 17, 25.

³⁸ Tindal and Fuchs, *Summary of Research*, 18, 24.

³⁹ Brooks, Case and Young, *Timed Versus Untimed*, 3 – 6. See Rindler, “Pitfalls”, 261 and Tindal and Fuchs, *Summary of Research*, 19 – 20, regarding quality of research.

⁴⁰ Brooks, Case and Young, *Timed Versus Untimed*, 3-4.

⁴¹ Brooks, Case and Young, *Timed Versus Untimed*, 7 – 9. Table 1 tells which skill sets were measured.

⁴² Brooks, Case and Young, *Timed Versus Untimed*, 7 – 9.

⁴³ Brooks, Case and Young, *Timed Versus Untimed*, 10, specifically:

This demonstrates that procedures established by Pearson over the past 80 years for setting time limits have allowed adequate time for non-disabled students to complete the test without undue strain or errors owing to time pressure. More important, however, is the evidence that the allowance of extended times accommodates disabled students so that they may demonstrate what they have learned while not unfairly inflating the scores of non-disabled students.

In fact, they argued that the extra time was sometimes counterproductive, with certain students who took the additional time to change correct answers. None of this argues against untimed testing for the LSAT, because it has already been established that the LSAT is partially speeded. There is not sufficient time for many of the test-takers to complete one answer to a given question, not to mention two.

A 2005 study by the College Board, owner of the SAT, surveyed the previous research and its own administration of the 1991 SAT, which was taken by 1.5 million students.⁴⁴ It found that “extra” time (1.5 or 2.0 times the standard or fixed time) for test-takers did tend to improve test results for high-achieving and medium-achievers, but not for low achievers.⁴⁵ Disabled students improved the most.⁴⁶ Improvement for math sections was significantly greater than for verbal sections. The researchers theorized that:

the verbal sections contained more items (nearly 60 percent) for which extra time will not be helpful (verbal analogy and sentence completion versus reading comprehension). Either the student knows the answer or does not, and extra time will not help. Extended time on the math items, in contrast, may allow the student the opportunity to work through the problems and obtain correct answers.⁴⁷

In the College Board study, performance for disabled and non-disabled students alike was better when section breaks were present than when students were allowed to pace themselves, with no section breaks.⁴⁸

Other studies, in which there was testing on college students and their peers not attending college, indicated that an untimed or extra-time test tended to increase performance versus a fixed-timed test, especially for disabled students.⁴⁹

There has been limited study into the question of whether additional instruction on test-taking skills would increase performance on standardized tests.⁵⁰ The studies reveal that such training does increase performance, though more research is necessary.⁵¹

College Board and Pearson conduct untimed and extra-time tests routinely, have developed protocols, and have published assessments.⁵² This demonstrates that such testing can be accomplished as a practical matter and without incident. In the Pearson case, they advise their proctors that the untimed Stanford Achievement Series test:⁵³

⁴⁴ Ellen B. Mandinach, Brent Bridgeman, Cara Cahalan-Laitusis, and Catherine Trapani, *The Impact of Extended Time on SAT® Test Performance: College Board Research Report No. 2005-8, ETS RR-05-20* (New York, NY: The College Board, 2005), <https://professionals.collegeboard.com/profdownload/pdf/CBR0508.pdf> (accessed approximately March 1, 2011).

⁴⁵ Mandinach, *et. al.*, *The Impact of Extended Time*, see, especially, pp. 1 – 5.

⁴⁶ Mandinach, *et. al.*, *The Impact of Extended Time*, 16.

⁴⁷ Mandinach, *et. al.*, *The Impact of Extended Time*, 17.

⁴⁸ Mandinach, *et. al.*, *The Impact of Extended Time*, 16.

⁴⁹ Tindal and Fuchs, *Summary of Research*, 19. Halla tested 126 undergraduate, graduate and non-students from ages 20 – 56 with the Graduate Record Examination (GRE) General Test and the Nelson-Denny Reading Test; “test scores increased significantly for both groups between timed and untimed testing conditions,” Tindal and Fuchs, *Summary of Research*, 22. Other studies of varying quality revealed that sometimes an untimed test improved scores, other times it did not. Tindal and Fuchs, *Summary of Research*, 21 – 6. There was only one study (through 2000) where increased time harmed performance. Tindal and Fuchs, *Summary of Research*, 19.

⁵⁰ Tindal and Fuchs, *Summary of Research*, 82 – 85.

⁵¹ Tindal and Fuchs, *Summary of Research*, 82 - 85.

⁵² See Case, *It’s About Time* and Mandinach, *et. al.*, *The Impact of Extended Time*.

⁵³ See Case, *It’s About Time* and Mandinach, *et. al.*, *The Impact of Extended Time*, 3 – 4.

- Typically will be completed by 95% of all students within 30 minutes of the fixed-time administration;
- Allows for the adoption of policies to turn the test from untimed to extra time, such as permitting students to work longer than 30 minutes beyond a guideline only if there are written instructions to that effect, for that student;
- Are still standardized tests and will be judged accordingly;
- Are designed to “give all students the opportunity to respond to each problem,” but some students will not “provide a response to every problem;”
- Etc.

4.6.4. LSAC’s analysis of “speededness” in the LSAT

LSAC has rarely looked into the question of whether test-takers are being rushed during the LSAT (i.e., is the test “speeded”) and, therefore, being denied an opportunity to demonstrate their true abilities. There seems to be little doubt that the LSAT is speeded.

In *Modeling Item Response Times With a Two-State Mixture Model: A New Approach to Measuring Speededness (CT 96-02)*, the researchers state:⁵⁴

Speededness refers to the extent to which time limits affect test takers’ performance. With regard to the Law School Admission Test (LSAT), speededness is currently measured by calculating the proportion of test takers who do not reach each item on the test. These proportions typically increase slightly toward the end of the test, indicating that the LSAT is partially speeded.

While acknowledging speededness, they also suggest that LSAC’s definition is inadequate:⁵⁵

Traditional speededness indices depend on “unreached” items, but when test takers engage in rapid-guessing behavior, the items are not unreached. More recent approaches to speededness depend on the drop in accuracy at the end of the test ... The latter approach is better because it incorporates guessing, but a more direct approach would focus on the amount of time spent on each item.

Their conclusions are stark. In a test similar to the LSAT, the researchers found that using a speededness measure of items-not-reached, 15% of the test-takers fell into this category and the test would have been considered speeded by traditional standards.⁵⁶ However, when a model that attempts to identify rapid guessing behavior is used, 50% of the test-takers exhibited this behavior, indicating much stronger evidence of speededness.⁵⁷

⁵⁴ Deborah L. Schnipke and David J. Scrams, *Modeling Item Response Times With a Two-State Mixture Model: A New Approach to Measuring Speededness, Computerized Testing Report 96-02* (Newton, PA: Law School Admission Council, March 1999), 2, <http://www.lsac.org/LSACResources/Research/CT/CT-96-02.pdf> (accessed approximately March 1, 2011).

⁵⁵ Schnipke and Scrams, *Modeling Item Response 96-02*, 15.

⁵⁶ Schnipke and Scrams, *Modeling Item Response 96-02*, 16.

⁵⁷ Schnipke and Scrams, *Modeling Item Response 96-02*, 16.

In *Comparison of LSAT Performance Among Selected Subgroups (SR 90-01)* at <http://www.lsac.org/LsacResources/Research/SR/SR-90-01.pdf>, the researchers state:⁵⁸

If general, a test section is not considered speeded if over 90 percent complete it. Over 90 percent of test takers from all subgroups except blacks and Puerto Ricans complete each section of the test. These data also confirm that the sections are differentially speeded by subgroup and that they are moderately speeded for Black/Afro-American and Puerto Rican test takers.

The researchers use the term “moderately speeded” because only 2.5% of the questions were either omitted or unanswered for both subgroups.⁵⁹ However, as noted above, a reasonable strategy for the LSAT is to guess, particularly when time has nearly expired. So, because rapid-guessing was never assessed by this study, the speededness observed by the researchers has almost certainly been underestimated.

The LSAC study *Speed as a Variable on the LSAT and Law School Exams (RR-03-03)* applies statistical techniques to actual LSAT scores, but does not examine the test itself. In this study, the researchers define part one of their study of a prominent national law school and a large regional law school as follows:⁶⁰

Student performance [in law school] is disaggregated into three distinct testing methods with varying degrees of time pressure: (1) in-class exams, (2) take-home exams, and (3) assigned papers. Correlation coefficients for both the LSAT and undergraduate grade-point average (UGPA) are then calculated. If test-taking speed affects both the LSAT and law school exams, the LSAT should be a relatively robust predictor of scores on in-class exams, with lower correlations on take-home exams and papers.

Unfortunately, this methodology is flawed because there could be multiple reasons for high or low correlations, and many variables for which the study does not control could be present. High correlations, for instance, could be expected between the LSAT and in-class exams, if **either** the LSAT and in-class exams are comparably speeded **or** comparably **non**-speeded. Throughout the study, the researcher speculates as to why the results appear to match or not match expectations. Unfortunately, there are always other reasonable explanations.⁶¹ One finding is that the LSAT consistently is better correlated

⁵⁸ Linda F. Wightman and David G. Muller, *Comparison of LSAT Performance Among Selected Subgroups, Statistical Report 90-01* (Newton, PA: Law School Admission Council, June 1999), 6, <http://www.lsac.org/LsacResources/Research/SR/SR-90-01.pdf> (accessed approximately March 1, 2011).

⁵⁹ Wightman and Muller, *Comparison of LSAT Performance*, 8.

⁶⁰ William D. Henderson, *Speed as a Variable on the LSAT and Law School Exams, Research Report 03-03* (Newton, PA: Law School Admission Council, February 2004), 1, <http://lsacnet.lsac.org/LSACResources/Research/RR/RR-03-03.pdf> (accessed approximately March 1, 2011).

⁶¹ Henderson, *Speed as a Variable*, 24. The author acknowledges this at the end of the study with: Notwithstanding any aggressive theorizing by the author, the findings of this study should be viewed as tentative and open to additional interpretation.

with in-class exams than either take-home exams or papers. This might not be due to speededness of both types of exams, but to other underlying similarity between all tests given in-class. Test-taking skills, test anxiety and rational vs. irrational responses to time pressure could all predict correlations between LSAT and in-class law school exam performance.

Though flawed, this study is an important beginning for LSAC and is valuable if we look to it as an indicator, not a certainty.⁶² The LSAC researchers reinforced the findings of the earlier studies, concluding that the LSAT was, in fact, speeded.⁶³ They also concluded that whether people of color are disproportionately impacted by this speededness was unclear.

4.6.5. Applicability of Research to the LSAT

The LSAT is designed to measure “acquired reading and verbal reasoning skills.”⁶⁴ The research, including the College Board SAT study of 1.5 million students and several LSAC studies, suggests that an untimed (or extra-time) LSAT would permit test-takers to demonstrate higher performance, because reading comprehension and verbal reasoning questions can frequently be solved correctly when the student has more time to work through the problem. The research also suggests that additional preparation with regard to test-taking skills would increase test scores.

There is evidence that law schools intentionally place students under time pressure on in-class exams and that a majority of law school courses have exams that are taken inside the classroom.⁶⁵ Some argue, therefore, that a speeded LSAT is an acceptable indicator of performance in a pressure-filled law school. Unfortunately, it is universally understood that standardized tests are designed to measure ability, not speed. So, arguing that a speeded LSAT is just a reflection of law school or anything else is contrary to the mission and purpose of a standardized test and LSAC’s own description of the purpose of the LSAT (recall, it is designed to measure skills, not speed).⁶⁶

However, it is possible to provide the speeded nature that proponents and apparently some law schools wish to see. This is partially what exists today, by the findings of LSAT’s own researchers. We can give a cleaner and more honest picture of speededness, as described below.

⁶² Where the study is particularly valuable, at least with respect to this communiqué, is the extent to which it summarizes other research on LSAT’s possible speededness and related topics.

⁶³ Henderson, *Speed as a Variable*, 1. The researchers concluded:

Overall, these results suggest that testing method is a variable that affects the ordinal ranking of law school performance and, therefore, the predictive power of the LSAT and UGPA. They are also consistent with the study’s hypothesis on test-taking speed.

⁶⁴ LSAC, *The Official LSAT SuperPrep*, 1.

⁶⁵ See Henderson, *Speed as a Variable*, 5 – 7.

⁶⁶ LSAC, *The Official LSAT SuperPrep*, 1.

4.7. Recommended LSAT Modifications

Based on the assumptions and research indicated above, the following recommendations are clear, practical and relevant. Note that there are other recommendations that do not involve LSAT modifications presented later.

Table 4: General Recommendations for Reforming LSAT, Based on Assumptions & Constraints

Assumption/Constraint	LSAT Modifications
1. Minimize test preparation inequities.	<p>1(a) Employ untimed or extra-time testing universally.</p> <p>1(b) LSAC will provide comprehensive, free and digestible characterizations of the test; what the test measures; the knowledgebase that needs to be mastered; all relevant test-taking skills.</p> <p>1(c) Eliminate questions and question types that do not add significantly to LSAT prediction capabilities.</p>
2. Consistently measure ability.	<p>2(a) Add a power score to supplement today's partial speed score. Possibly state a minimum number of questions to be attempted to qualify for the calculation of the power score.</p> <p>2(b) LSAC will provide guidance explaining power score as an alternative to demonstrate ability without preparation expense, drills, sacrifices. Encourage students to choose one of the two scores as their target for that exam: power or speed.</p> <p>2(c) Law schools can set admission criteria based on one or both of the scores.</p>
3. Retain intrinsic quality of differentiating among the best prepared.	<p>3(a) Continue to provide current LSAT score so those who benefit from today's inequitable system may continue to compete against each other.</p> <p>3(b) Alternatively, enable existing LSAT score to generate more competition by increasing number of questions in each section. Call this a speed score. This will enhance competitive balance among best prepared and maximally reward those who want no changes to current system. This increases the likelihood that the power score will be used primarily, though not exclusively, by those with high LSAT test-taker's preparation burdens.</p> <p>3(b) LSAC will provide guidance on tradeoffs between power and speed scores. Adding significantly more questions will make it impossible to maximize both.</p>

In the following sections, we describe each of the above modifications and their implications.

4.7.1. Modification Set #1: To minimize test preparation inequities

4.7.1.1. *1(a) Employ untimed or extra-time testing universally*

We have provided evidence that LSAC is aware that the LSAT is partially speeded and, hence, does not measure ability with accuracy. The College Board has demonstrated that untimed tests or extra-time tests are likely to improve test scores of those who possess the ability to solve the problems on the test, but who simply lack the time. The LSAT questions are not those of vocabulary or word associations, which the test-taker either knows or does not know. Instead, the LSAT contains problems that can be solved by a test-taker who possesses the ability to solve them – but only when there is sufficient time.

Pearson and College Board now administer untimed and extra-time testing with little challenges, so this modification is practical.

There is an argument that speeded testing mirrors high-pressure law school exams and legal practice, so the LSAT should remain speeded. However, an LSAT in which significant portions of the test-takers cannot read all of the questions and must rapidly guess because the time limit is about to expire is a depiction contrary to LSAC's assertions that the LSAT measures ability, not speed. And, even if some law school in-class exams rush students, there is no evidence that this is the dominant model and that mere time-pressures necessarily deprive a student of the opportunity to demonstrate ability. (Recall that in a speeded exam, by contrast, test-takers do not have enough time to read and answer all of the questions.) As well, law school courses frequently require take-home exams and papers, not just in-class exams. Finally, there is no consensus on the question of whether a speeded LSAT or in-class exams prepare or test anyone for the rigors of legal practice.

The evidence, therefore, does not support any benefits for a speeded LSAT.

However, for those who believe a speeded LSAT test helps to separate the wheat from the chaff, when we have argued that the LSAT actually rewards those with a lower LSAT test-taker's preparation burden, LSAC can accommodate them as well. The number of test questions could be dramatically increased to a number that no one could possibly complete. The traditional LSAT score would then measure speed more than power and those with low LSAT test-taker's preparation burdens can compete against each other to demonstrate their preparedness. This would, essentially, be a "fair fight," because those with high LSAT test-taker's preparation burdens could be expected to concentrate on their power score. The traditional LSAT score could be renamed to a "speed" or "intensity" score.

4.7.1.2. *1(b) LSAC will provide comprehensive disclosures*

If the LSAT is to "measure skills considered essential for success in law school," and not force test-takers to spend valuable time and money seeking interpretations of that clause, LSAC should be required to share its internal documentation to describe, in excruciating detail, precisely what skills the LSAT is measuring, with endless examples. It should make all of this information available on its website. It should include recommended

test-taking strategies, study schedules and means of self-instruction. LSAC should be responsible for de-mystifying the LSAT so that everyone, not just the comparatively advantaged, will know how to maximize their chances to do well. In an untimed or extra-time administration, the typical test-taker concentrating on her power score will simply be asked to demonstrate what she knows, not beat the clock. She will be able to go into such a test confident that her abilities will open or close doors for her.

4.7.1.3. *1(c) Eliminate questions and question types that do not add value.*

Removing redundant question types or sections is one of the clearest and defensible ways of modifying the LSAT to reduce inequities associated with the LSAT test-taker's preparation burden and reduce test anxiety. It also produces other benefits.

4.7.1.3.1. *LSAC research demonstrates that AR and RC sections add redundancies*

LSAC's evidence regarding the predictive capabilities of the Analytical Reasoning (AR) and Reading Comprehension (RC) sections of the LSAT reveals that there is no additional benefit to retaining the AR and RC sections versus the Logical Reasoning (LR) section alone.

The LSAC's last published study of the predictive power of current LSAT question types is entitled *LSAT Item-Type Validity Study (TR 98-01)*. It can be found on the LSAC website at <http://www.lsac.org/LsacResources/Research/TR/TR-98-01.pdf>. The study compared how well each LSAT section (for 13 different LSAT tests from 1993 - 96) predicted first-year law school GPA (FYA), through the statistic "r" – the correlation coefficient.⁶⁷ Although not presented in the study, we calculate below the much more informative coefficient of determination, R^2 , which tells us precisely how much of the variance (essentially, fluctuation) in FYA is explained by LSAT scores.⁶⁸ We will discuss R^2 in detail later.

The study calculated different correlations (r) for all 168 law schools analyzed and reported only the average (mean) of the correlations. These results are summarized below.

⁶⁷ See MathBits.com, "Correlation Coefficient: How well does your regression equation truly represent your set of data?," <http://mathbits.com/mathbits/tisection/statistics2/correlation.htm> (accessed approximately March 1, 2011). The quantity r, called the linear correlation coefficient, measures the strength and the direction of a linear relationship between two variables (e.g., LSAT score and FYA). The value of r is such that $-1 < r < +1$. If x and y have a strong positive linear correlation, r is close to +1. Positive values indicate a relationship between x and y variables such that as values for x increases, values for y also increase. Negative values indicate a relationship between x and y such that as values for x increase, values for y decrease."

⁶⁸ MathBits, "Correlation Coefficient." The coefficient of determination, R^2 , gives the proportion of the variance (fluctuation) of one variable (e.g., FYA) that is predictable from the other variable (e.g., LSAT scores).). The value of R^2 is such that $0 < R^2 < 1$. It is a measure that allows us to determine how certain one can be in making predictions from a certain model/graph. The coefficient of determination is the ratio of the explained variation to the total variation. It denotes the strength of the linear association between x and y. It represents the percent of the data that is the closest to the line of best fit.

Table 5: Correlations between LSAT Sections and FYA
from *LSAT Item-Type Validity Study (TR 98-01)*

Item	Mean “r”: LSAT scores vs. FYA*	Mean R ² : LSAT scores vs. FYA*
AR	0.244	0.059
LR	0.358	0.128
RC	0.314	0.098
Total Scaled Score	0.388	0.150

*FYA is an abbreviation for first-year law school GPA.

LSAC study measured results for 13 test administrations and 168 law schools.

Clearly, the total scaled score correlation is only marginally better than the correlation for the LR section alone. **From R², we know that the total LSAT score accounts for only 2.2% more of the fluctuation in FYA than the LR section score.**⁶⁹ Equivalently, predictions of FYA from the total LSAT score explain only 2.2% more of the variability of FYA about the regression line than a prediction made with the LR section score alone. Assuming that additional predictability for the total score is, roughly, evenly distributed among all 61 possible LSAT scores, it is little more than rounding error.⁷⁰

Later in the above-referenced study, there is an attempt to perform a statistical transformation on the above data to account for the so-called “restriction of range” problem.⁷¹ Unfortunately, this transformation still yields only an approximation, with one examination of its application finding that it over-corrected the “actual” correlation by 14%.⁷² Even this transformation, however, reveals that use of the LR score alone

⁶⁹ The R² difference between the *LR correlation with FYA* and the *entire LSAT score correlation with FYA* was 0.022 (0.150 - 0.128). Therefore, considering the R² values, the total LSAT score predicts the FYA with an accuracy that accounts for only 2.2% more of the variation in FYA than a prediction made with the LR section score alone.

⁷⁰ Some might argue that 3.1% is a large number because the 25-percentile to 75-percentile LSAT scores of admitted students in top law schools equates to roughly the top 1 percent of LSAT scores. However, the LSAT data provides no support for the notion that such predictability would be concentrated at the top end of the LSAT spectrum. One would have to assume, therefore, that this 3.1% additional ability to explain results is spread among all possible 61 LSAT scores. AR and RC scores, therefore, add predictability to any particular LSAT score to a negligible, perhaps non-quantifiable (if margins of error are factored in) extent. Additional study would clarify this, but the difference is so negligible, very little insight would likely be gained.

⁷¹ Marie Wiberg and Anna Sundström, “A Comparison of Two Approaches to Correction of Restriction of Range in Correlation Analysis,” *Practical Assessment, Research & Evaluation* 14, 5 (March 2009), <http://pareonline.net/pdf/v14n5.pdf> (accessed approximately March 1, 2011).

⁷² Wiberg and Sundström, “A Comparison”, 6. It overstated the correction when direction is taken into account and understated it when only magnitude is considered. The standard deviations were 3 – 6 times greater in this Swedish study than in the LSAC study, however, so directly comparing the magnitude of the corrections is not helpful.

produces predictability results that are almost identical to those yielded by the entire LSAT score.⁷³

Hence, based on the 13 LSAT test administrations measured, the LR section predicts FYA as well (or poorly) as the total LSAT score.

4.7.1.3.2. *Continued inclusion of AR and RC wastes resources*

Since they add nothing beneficial to the predictive aspects of the admissions process, their continued use is superfluous. There is a waste of resources associated with maintaining, continually updating and scoring three different sections of the LSAT. More focus on LR by the LSAC could refine the questions and easily produce the additional test items necessary for an expanded LR.

4.7.1.3.3. *Removal of AR and RC would decrease the intense anxiety associated with LSAT*

A typical LSAT test trainer and any of the commercial test-preparation schools will confirm that students are initially mortified when they encounter the AR section. It is only through intense instruction and/or self-study that they become more comfortable. Confirmation of the terror initially faced by students with regard to the AR section can be found by looking at any of the legal blogs. One can quickly see that the AR section is the one where preparation resources (time and money) as well as anxiety create immediate inequities in the LSAT preparation process.

Though more study would be helpful, it seems abundantly clear that removing the AR section would alleviate much of the financial, time and psychological imbalances of the LSAT. And, it would not harm the predictability of the LSAT. This is an easy win-win for all concerned.

With respect to the RC section, blogs and trainers report anxiety in preparation for it, as well, particularly among slower readers. PowerScore states:

Unlike most of the reading that you have done in your daily life (magazines, newspapers, novels, etc.), Reading Comprehension passages are not designed to present ideas in the clearest way possible. Instead, the test makers do just the opposite: they purposely make the passages complicated and confusing.

Preparation for such questions can be painstaking. More independent research would be helpful, but AR and RC preparation would appear to be significant and would necessarily place those who lack time and resources at a decided disadvantage versus those test-takers who have no such limitations.

⁷³ The revised values provided by the study were $r = 0.483$ for *LR v. FYA*, and $r = 0.529$ for *total LSAT raw score v. FYA*. The R^2 difference between the *LR correlation with FYA* and the *entire LSAT raw score correlation with FYA* was only 0.047 ($= 0.280 - 0.233$). **This means that total raw score predicts FYA by accounting for only 4.7% more of FYA's variance than the LR section score alone.** As in the uncorrected case, even this "corrected" calculation demonstrates that the entire LSAT score is only negligibly better than the LR score alone.

4.7.1.3.4. Removal of AR and RC would decrease the inequities in student preparation for the LSAT

It is axiomatic that having to prepare for three sections is a far more challenging process than having to prepare for just one. Dropping AR and RC would allow any prospective student with a high LSAT test-taker's preparation burden to have a more affordable, manageable preparation task and would require her to make fewer personal sacrifices.

4.7.1.3.5. Removal of AR and RC would decrease racial, ethnic and gender biases present in the LSAT

The AR and RC sections exhibit the most racial, ethnic and gender spread of the three sections. ***Since the RC and AR sections do not add to the predictive value of the LSAT, and in their absence there is less racial, ethnic and gender differentiation in test scores, their presence appears to be, in part, predicting differences in race, ethnicity and gender unrelated to law school success.*** Thus, removing them would have the desired policy outcome of removing predictors that might be measuring racial/ethnicity differences more than law school-relevant data.

The latest LSAC study on how those of different race/ethnicities perform on individual sections appears to be *Comparison of LSAT Performance Among Selected Subgroups (SR 90-01)* at <http://www.lsac.org/LsacResources/Research/SR/SR-90-01.pdf>. Table 5 of this study is reproduced below.

Table 6: Percent Answered Correctly for Each Section, by Race, Ethnicity & Gender
from *Comparison of LSAT Performance Among Selected Subgroups (SR 90-01)*

Subgroup	Analytical Reasoning Section	Reading Comprehension Section	Logical Reasoning Section
Gender			
Male	62%	63%	67%
Female	62%	59%	66%
Race/Ethnicity			
Asian/Pacific Islander	65%	60%	64%
Black	45%	49%	53%
White	64%	63%	68%
Hispanic	55%	56%	60%
Puerto Rican	49%	51%	55%
Mexican American	52%	55%	59%

Note that this comparison indicates that RC differentiates most with respect to gender and, given the duplicative nature of the RC section versus LR-only, those measured differences would appear to have nothing to do with the assessment of ability or aptitude measured by LR's correlations (presented earlier). Removing the RC section, therefore, will also remove this unwanted differentiation.

Again, more study would be helpful. But, the key message is that there would likely be only benefit, nothing lost, if the AR and RC sections were removed.

4.7.2. Modification Set #2: To consistently measure ability, introduce a power score

As noted above, the LSAC has defined the mission of the LSAT as “to measure skills considered essential for success in law school.” Today’s partially speeded LSAT cannot, by definition do this. Inequities resulting from a high LSAT test-taker’s preparation burden will be alleviated by the measures indicated earlier. Introducing a power score (one that measures accuracy) takes this to the next level.

4.7.2.1. *The Mastery score, what it would assess, how it would be used*

Per the preceding section, prospective law students would have ample time to wade through virtually any problem on the LSAT, without the benefit of inequitable test preparation courses, materials, etc., if there were sufficient time allocated for each section. A measure, then, of consistent assessment of logical reasoning and reading comprehension ability is a simple power score. For clarity, we will call it a ***Mastery***, score and it would be obtained by computing the percent answered correctly on an *untimed* or an extra time test. If only a fixed-time LSAT is available, then we calculate it as the ***percent answered correctly for whatever number of questions the test-taker completes***, subject to a comfortable minimum of perhaps 10-15 questions answered per section.

This target minimum would be extremely low for a law school that wanted to measure the fundamental logical reasoning ability of the most disadvantaged students – who presumably had neither the time nor resources to undertake significant preparation. Prospective students could proceed at a pace comfortable to them, with the full knowledge that their reading speed and ignorance of test-taking maxims and shortcuts will not necessarily prevent them from demonstrating competency.

The Mastery score, then, consists of two parts:

- a. The number of questions attempted.
- b. The percent answered correctly.

Naturally, law schools can designate their own baseline to achieve their desired diversity objectives, while not sacrificing what they deem to be competency. At the other end of the spectrum, law schools could finely tune their selectivity using this measure by favoring applicants who have completed a greater number of questions and achieved a higher percentage correctly answered.

4.7.3. Modification Set #3: To enhance differentiation among the best prepared, introduce a speed score

Complementary to the Mastery score, which is a pure measure of ability designed to alleviate test inequities, is a scoring measure that will allow the LSAT to retain its capacity to differentiate among those test-takers who had the time and money to prepare maximally for the LSAT.

4.7.3.1. *The Intensity score, what it measures, how it would be used*

To differentiate among the students who had the greatest advantages of time and resources devoted to LSAT preparation (including, perhaps, commitment), the equivalent of today's highly structured, pressure-filled time trial should be maintained. Keeping the current scoring, in addition to the new Mastery score, will accomplish that, but only if the test remains speeded, that is, the LSAT remains a timed test as it is today, with no extra time permitted.

Should the LSAT become an untimed test, we would expect that the highly prepared would have sufficient time to answer all questions correctly. The current scoring would no longer be speeded and we would have essentially two Mastery scores: one with the advantages of significant preparation, one for those with significant LSAT test-taker preparation burdens.

Should the LSAT become an extra-time exam, then there will likely be sufficient opportunity, as in the untimed case, for the very best prepared to answer every question correctly and the current scoring would simply become a second Mastery score for the advantaged and extraordinarily committed. However, unlike the untimed case, the extra-time exam can be modified to reproduce that speeded nature that seems to be desirable among elite universities and the advantaged.

To accomplish this, the total number of questions should be increased to an amount that no one could possibly answer, even with the extra time added – perhaps 50 per section. The current scoring measure would then be returned to a speeded state. Of course, this would be highly speeded, much more so than today's LSAT. This score, which we can call an **Intensity** score, will measure in the manner of today's score:

- Score equals the number answered correctly;
- There will be no penalty for guessing.

This score is explicitly demonstrative of intense preparation, with time sensitivities amplified. An increased number of questions on each section would guarantee distinctiveness to each type of score – the more questions one attempts in an effort to raise his Intensity score the more he reduces his opportunity for a high Mastery score. All but the exceptionally talented would have to choose which score they wished to maximize.

4.7.3.2. *Law school application of Intensity scores*

Law schools that wish to identify those best prepared for the LSAT, in some belief that they will be better students through privilege, commitment, skill, preparedness, etc., can over-emphasize the Intensity score and de-emphasize the Mastery score or greatly emphasize both. Differentiating between exceptional minds and the test-takers fortunate enough to obtain high-quality preparation would be laid bare, through competition for the highest Intensity scores. On the other hand, though poor quality preparation might not stop a student from demonstrating Mastery, it should significantly harm the Intensity score. All law schools could choose the mix of scores that would allow them to throttle

diversity based on logical reasoning capacity, ability to prepare for the test, and absolute mastery of the skill.

It would likely be a frequent occurrence that a bright, struggling single parent, desperately putting her affairs in order before law school the following year, but who has little resources to prepare at the level of the most advantaged competitors, will obtain a Mastery score superior to said competitors. It would be as frequent an occurrence, of course, that her Intensity score will be inferior.

4.7.4. Combining both Mastery and Intensity scores

4.7.4.1. *Additional data points can be gathered*

At the commencement of any current LSAT test administration, test takers are asked to answer several questions, some optional. At the end of the test, the test-taker is given an opportunity to cancel his or her test results. For the new LSAT, several additional questions could be asked to assist with interpretation of test results and to help with future administrations. Such questions include the following:

- Whether the test-taker is seeking primarily a Mastery score, an Intensity score or both. This would allow the test-taker to explain his performance, should the question arise. It might also be useful to admissions officers – if the answers were released to them.
- How much preparation time and expense the test-taker devoted and the nature of the preparation efforts. This would demonstrate some combination of commitment, privilege and opportunity. This would be helpful to studies of the test results. This information might also be helpful to admissions officers – if the answers were released to them.
- Whether the test-taker had to make sacrifices to prepare for the exam and the nature of those sacrifices.
- The general state of mind of the test-taker during the preparation period: anxiety level, impact on work, school, family.
- Whether the test-taker would like to participate in any follow-up survey or conversation about the exam.
- Whether the test-taker wished the answers to these preliminary/post-test questions to be released to law schools or any other entities.

If answering these questions would be considered too distracting at the beginning of the exam, then they could be asked at the end of the exam or could be completed prior to entering the test site. Perspective would be provided if the administrator introduced the 3-5 minute questionnaire as an opportunity to improve the LSAT and repay those whose answers in previous test administrations have already helped that day's test-takers.

4.7.4.2. *Avoiding the stigmatization of Mastery scores*

Care must be taken to avoid any possible stigmatization of the Mastery score. If this is seen as an inferior path for unqualified people to be admitted to law school, then it will be short-lived. In this regard, it would be helpful to publicize the fact that those who take the LSAT, *solely* for their Mastery scores, regularly achieve higher such scores than those with Intensity scores of 170+.

We assume that law schools will allow a certain portion of their available seats to be filled by using Mastery scores, in lieu of Intensity scores. However, Mastery scores would become increasingly divisive if race, ethnicity and/or gender were used as criteria for weighing Mastery scores more heavily in admissions decisions. Constitutional questions would also arise at some point in the near future, as discussed below. Care should be taken to offer the disproportionate weighting of Mastery scores to those who have greater LSAT test-taker preparation burdens, or those who bring something special and rewarding to the educational environment. Within this context, the judicious use of Mastery scores and de-emphasis of Intensity scores can serve a number of lawful, desirable purposes. The following are examples:

- Law schools could limit Mastery score admits to those who have proven, extraordinary academic and other abilities and for which the LSAT would be an unnecessary burden: those already holding advanced degrees, artists, those who request it as part of an early admissions program, applicants who have some exceptional ability that tends to be masked by standardized tests, etc.
- Certain law schools could limit Mastery score admits to those likely to diversify the campus and classroom (avoiding direct racial, ethnicity and gender-specific criteria): those above a certain age, those within the top 10% of the GPA's at an HBCU or any "non-mainstream" school, applicants within certain undergraduate GPA ranges, etc.
- Certain law schools could limit Mastery score admits to those likely to be *unable* to prepare for the Intensity score component due to hardship or other reasons: those above a certain age, those within certain family income ranges, applicants with significant family responsibilities, those with disabilities, etc.

Alternatively, some law schools could choose to favor neither score, but will select the highest of the two scores for all applicants, possibly using the other to break ties. This avoids any possible stigmatization of Mastery scores. And, in extra-time, LR-only LSAT administrations, where full disclosure of what the test attempts to measure has occurred, LSAT inequities could still diminish significantly and diversity might still prevail. This is easily the most desirable outcome, but this requires careful study.

4.7.4.3. *Importance of defense measures that would lessen or eliminate opportunities to exploit the new regime*

To eliminate gamesmanship, law schools could enact appropriate disincentives. These include the following:

- Law schools could choose to weigh only Mastery and Intensity scores from the *same* LSAT exam. This would eliminate the scenario in which one would take the new LSAT twice, once to excel at the Mastery score and once to excel at the Intensity score. Each prospective student would have to choose which measure they thought best suited their test preparation strategy and admissions goals.
- Law schools could ensure that the percentage of seats in their entering class targeted to be filled based, in part, on Mastery scores would be significantly smaller than those filled primarily based on Intensity scores. This would make the more advantaged prospects less likely to attempt to qualify on the basis of a Mastery score.

- To increase the likelihood that the disincentives will be effective, law schools should publish not only the disincentives, but precisely how they weigh each score. There would be nothing to hide. These are 100% merit-based measures.

4.7.4.4. *Mastery scores and LR-only testing need each other*

Having only one kind of test item reduces the complexity and preparation inequities that will arise in any new test. So the Mastery score strategy could most equitably be applied in an LR-only world. Attempting to apply it to today's exam would still force the poor and disadvantaged to learn multiple question types (particularly AR), correspondingly amplify their anxiety and, invariably, demonstrate an inability to match or even approximate their more advantaged competitors in terms of time and resources. They would have a somewhat clearer path to demonstrating fitness for law school, but their burdens would still be significant and the test results in each section would show this.

4.7.4.5. *Assumptions must be validated with independent studies*

More study of the Mastery score and other creative measurement devices is, of course, needed. In fact, if studies do not confirm the underlying premises indicated above, then there is little basis for several of the modifications presented above. Specifically, independent studies need to confirm two assertions. First, there is a correlation between LSAT test-taker's preparation burden (time, resources and sacrifices required) and LSAT score. Second, the LR section predicts FYA with a correlation comparable to the entire LSAT score. These two correlations should be at approximately the levels reported by LSAC for its LSAT vs. FYA correlations: $r \approx 0.30$. (We argue in this communiqué that such correlations are extremely weak, but it seems fitting to accept LSAC's standard when evaluating criticisms of the LSAT.)

5. PROPOSAL B. LSAC’S STUDIES CONTAIN ERRORS, PECULIAR CHOICES, SPIN THAT DISTORTS THE POOR PERFORMANCE OF THE LSAT, AND A CLEAR REFUSAL TO INVESTIGATE QUESTIONS DAMAGING TO THE LSAT.

The LSAC studies that purport to apply scientifically sound methodology and analysis to determine the validity of the LSAT and to address various LSAT controversies contain errors, inexplicable methodological choices, misstatements and omissions that undermine LSAC’s credibility. Accrediting agencies must re-evaluate every LSAC study and any other basis for mandating LSAT testing for law schools they accredit. Law schools must perform independent research to determine whether any continued reliance on the LSAT is warranted.

Of particular note in the sections that follow is a discussion of the correlation coefficient (r) versus the coefficient of determination (R^2). We explore the mathematics and psychology of LSAC’s use of r , in lieu of the more intuitive and meaningful R^2 . We re-interpret two decades of LSAC studies through the latter measure and demonstrate how damaging the faulty reliance on the LSAT by admissions officers and accreditation bodies has been. We discuss the urgent need for reform.

5.1. Errors and peculiar choices in studies

This section has been moved to Exhibit A, for readability.

5.2. Spin that distorts the poor performance of the LSAT

5.2.1. The pervasive, deceptive and inappropriate use of the correlation coefficient (r) and refusal to use the coefficient of determination (R^2)

5.2.1.1. *Introduction*

Consistently, throughout virtually every one of LSAC studies, the researchers rely upon the correlation coefficient, r , and attribute a level of validity to it that the mathematics, literature and even scatter plots indicate is utterly unjustified. Though we will discuss a more intuitive sense of correlation in detail below, so that the reader can make an informed decision regarding the levels being measured by the LSAT and their appropriate applicability to law school admissions, the most generous view of the correlation coefficient is that values between 0.3 and 0.5 are considered weak or moderately correlated.⁷⁴ Virtually every LSAC calculation of r is between 0.2 and 0.4.

⁷⁴ For instance, MathBits, “Correlation Coefficient,” states that “A correlation greater than 0.8 is generally described as *strong*, whereas a correlation less than 0.5 is generally described as *weak*.” Jacob Cohen describes the correlations measured at $r > 0.5$ as “large” in social sciences and $r = 0.3$ as medium. See Jacob Cohen, *Statistical power analysis for the behavioral sciences, 2nd Edition* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1988), 78-81, [http://books.google.com/books?hl=en&lr=&id=TI0N2IRA09oC&oi=fnd&pg=PR11&dq=Cohen,+J.+\(1988\).+Statistical+power+analysis+for+the+behavioral+sciences+\(2nd+ed.\)](http://books.google.com/books?hl=en&lr=&id=TI0N2IRA09oC&oi=fnd&pg=PR11&dq=Cohen,+J.+(1988).+Statistical+power+analysis+for+the+behavioral+sciences+(2nd+ed.)) (accessed approximately March 1, 2011). He cautions against considering this a firm rule and his reasons, described elsewhere in this communiqué, are not helpful to the LSAT.

5.2.1.2. *Correlation coefficient (r) as “proportion of variance” fails as an intuitive measure*

There is one view of r that proponents argue is intuitive, the proportion of variation (*not variance*) in Y (e.g., FYA) accounted for by variations in X (e.g. LSAT score, LR score, undergraduate GPA).⁷⁵ One major problem with this interpretation is that the **proportion of variation NOT accounted for** should, intuitively, be $1 - r$. Instead, it is $\sqrt{1 - r^2}$, which is called the coefficient of alienation.⁷⁶ That the proportion of variation *accounted for* + *not accounted for* does not equal 1 is unsettling. Example, if $r = 0.40$ = the **proportion of variation accounted for**, then the proportion of variation *NOT* accounted for is **0.92**. There is nothing intuitive about that. But, since LSAC has decided to employ r, we are entitled to catalog what it does *not* explain, and do so in several tables below.

5.2.1.3. *The coefficient of determination, R², a standard for evaluating correlation*

For a quantitative and intuitive sense of how poor LSAC’s correlations really are, we employ the square of the multiple correlation coefficient (R); this value, R², is called the coefficient of determination, as described above.⁷⁷ Unlike r, R² provides a direct and intuitive measure of correlation and, fortunately, the relationship between r and R² is simple for the 2-variable prediction case found in virtually every LSAC study: $R^2 = r^2$.⁷⁸

⁷⁵ See Roy D’Andrade and Jon Dart, “The Interpretation of r Versus r² or Why Percent of Variance Accounted for Is a Poor Measure of Size of Effect,” *Journal of Quantitative Anthropology* 2 (1990), 52, [http://www.quantitativeanthropology.org/index.php?journal=QA&page=article&op=viewFile&path\[\]=28&path\[\]=44](http://www.quantitativeanthropology.org/index.php?journal=QA&page=article&op=viewFile&path[]=28&path[]=44) (accessed approximately March 1, 2011).

⁷⁶ D’Andrade and Dart, “The Interpretation of r.” See also McGraw-Hill Science & Technology Dictionary on answers.com, “Coefficient of Alienation,” <http://www.answers.com/topic/coefficient-of-alienation> (accessed approximately March 1, 2011). See also Robert M. Kaplan and Dennis P. Saccuzzo, *Psychological Testing: Principles, Applications, and Issues (7th Edition)*, (Belmont, CA: Wadsworth, 2008), 85. <http://books.google.com/books?id=4TieNj-MlrsC&pg=PA85&lpg=PA85> (accessed approximately March 1, 2011).

⁷⁷ See also StatSoft, “Multiple Regression,” *StatSoft Electronic Statistics Textbook* (Tulsa, OK: StatSoft, Inc., 2011), <http://www.statsoft.com/textbook/multiple-regression/#cinterpreting> (accessed approximately March 1, 2011) and MathBits, “Correlation Coefficient.” The coefficient of determination gives the proportion of the variance (fluctuation) of a predicted variable (e.g., FYA) that is predictable from the predictor variable (e.g., LSAT scores). From MathBits, “Correlation Coefficient”:

The *coefficient of determination* represents the percent of the data that is the closest to the line of best fit. ...

The *coefficient of determination* is a measure of how well the regression line represents the data. If the regression line passes exactly through every point on the scatter plot, it would be able to explain all of the variation. The further the line is away from the points, the less it is able to explain.

The *coefficient of determination* is such that $0 < R^2 < 1$, and denotes the strength of the linear association between x and y.

⁷⁸ The general formula for R², when y_i are observable values and f_i (sometimes called \hat{y}_i) are predicted, is (from Wikipedia contributors, “Coefficient of determination,” *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/w/index.php?title=Coefficient_of_determination&oldid=415871873 (accessed approximately March 1, 2011)):

$$R^2 \equiv 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}}, \quad \text{where} \quad SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2, \quad = \text{“total sum of squares”}$$

R^2 measures on a percentage basis, how well the LSAT (or its components) explain FYA. For example, in a case where LSAC asserts that a law school reports an r of 0.3 for LSAT's correlation with FYA, $R^2 = .09$ and this signifies that the LSAT only explains 9% of the variance in FYA.⁷⁹ There are special cases which make R^2 less useful, but they do not appear to be present here – and we leave it to LSAC to make that argument.⁸⁰

$$SS_{\text{err}} = \sum_i (y_i - f_i)^2 \quad \bar{y} = \frac{1}{n} \sum_i y_i$$

= “residual sum of squares”, = “mean of observed data”

$$SS_{\text{reg}} = \sum_i (f_i - \bar{y})^2,$$

= “regression sum of squares”

In a general form, R^2 can be seen to be related to the unexplained variance, since the second term compares the unexplained variance (variance of the model's errors) with the total variance (of the data). Therefore, $(1 - SS_{\text{err}}/SS_{\text{tot}}) = \text{proportion of explained variance} = \text{portion of variance attributable to model}$.

We can be explicit (see Wikipedia contributors, "Coefficient of determination") about the conditions under which $R^2 = r^2$:

In case of a single regressor, fitted by least squares, R^2 is the square of the Pearson product-moment correlation coefficient relating the regressor and the response variable. More generally, R^2 is the square of the correlation between the constructed predictor and the response variable.

See also Hedge Fund Contributors, “R-Squared,” *Hedge Fund Consistency Index*, http://www.hedgefund-index.com/d_rsquared.asp (accessed approximately March 1, 2011), Section 3, when a single predictor variable (such as the LSAT score) is used to predict or estimate a single variable (such as FYA) with ordinary least squares regression, $R^2 = r^2$. From Princeton University, “Interpreting Regression Output,” *Data and Statistical Services*, http://dss.princeton.edu/online_help/analysis/interpreting_regression.htm (accessed approximately March 1, 2011), this relationship is described as:

The R-squared of the regression is the fraction of the variation in your dependent variable that is accounted for (or predicted by) your independent variables. (In regression with a single independent variable, it is the same as the square of the correlation between your dependent and independent variable.) The R-squared is generally of secondary importance, unless your main concern is using the regression equation to make accurate predictions.

Finally, LSAC studies are simple, linear regression techniques. See Lisa Anthony Stilwell, Susan P. Dalessandro and Lynda M. Reese, *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*, *LSAT Technical Report 09-03* (Newton, PA: Law School Admission Council, October 2009), 3, www.lsac.org/LSACResources/Research/TR/TR-09-03.pdf (accessed approximately March 1, 2011). Therefore, for all analysis involving a single predictor variable, such as LSAT score, AR score, RC score, LR score, undergraduate GPA, $R^2 = r^2$.

⁷⁹ An example of how to interpret this measure is found in StatSoft, “Multiple Regression”:

If we have an R-square of 0.4 then we know that the variability of the Y values around the regression line is 1-0.4 times the original variance; in other words we have explained 40% of the original variability, and are left with 60% residual variability. Ideally, we would like to explain most if not all of the original variability. The R-square value is an indicator of how well the model fits the data (e.g., an R-square close to 1.0 indicates that we have accounted for almost all of the variability with the variables specified in the model).

⁸⁰ There are examples where a researcher states that R^2 does not adequately capture correlation. Those claims are sometimes easy to demonstrate, and, therefore, address. But, sometimes, they tend to be subjective and generally outside the mainstream of academic thought and practical experience. See, for example, the final pair of scatter plots where the author's view is not necessarily scientific, in Wolfgang A. Rolke, “Coefficient of Determination and Overfitting,” *ESMA 3102 Introduction to Statistics II*, <http://math.uprm.edu/~wrolke/esma3102/overfit.htm> (accessed approximately March 1, 2011). See also

5.2.1.4. *Correlation coefficient (r) is misleading and inferior to coefficient of determination (R²)*

The LSAC's publication of the correlation coefficient (r) and never the coefficient of determination (R²) gives at least two false impressions that promote a deceptive view of the LSAT that is beneficial to LSAC.⁸¹

First, some readers, administrators, policymakers, perhaps even accreditation bodies, will implicitly accept the interpretation of r as a percent or proportion of correlation. There is nothing in the literature to support this and it is simply untrue. There is a view in the literature that r can serve as a *proportion of variation*. While technically true, as demonstrated above, there is nothing intuitive about the concept, due to the failures of its complementary coefficient of alienation to behave intuitively and by the mathematics of r vs. R². The tantalizing prospect, however, that a correlation of 0.3 (R² → 9%) will be seen, falsely, by anyone as a 30% correlation can only help the perception that the LSAT is indispensable. This would appear to be a powerful inducement for LSAC to publish r values and avoid publishing R² figures.

Second, this false impression of r's interpretation as a percent of correlation is worsened by a numerical anomaly. A positive fraction between 0 and 1 has a square root that is always **larger** than the fraction itself. Because correlations fit this description, the magnitude of r is always larger than r². At the levels of correlation measured by the

several specialized examples purporting to show r being more intuitive and insightful than R²: D'Andrade and Dart, "The Interpretation of r." Of course, because, in the 2-variable case, R² = r² or |r| = sqrt(R²), the LSAC's correlations would be roughly as distorted as anything gleaned from R². So, we are not imposing an unfair measure of correlation, just using a variation of the one chosen by LSAC.

Nonetheless, we note the following. (a) When complex transformations arise which incorporate variance into the "known" variable, the R² value is distorted by the transformation and has diminished value. This does not occur in LSAC's simple linear regressions or their attempted restriction of range corrections. See Robert F. Nau, "What's a good value for R-squared?," at <http://www.duke.edu/~rnau/rsquared.htm> (accessed approximately March 1, 2011). (b) R² will decrease when there are a larger number of outliers, slope of the regression line flattens (FYA changes little with LSAT score), or when there is less range in the predictor variables for the same amount of scatter. See Colorado State University's Starmap Learning Materials at <http://www.stat.colostate.edu/~nsu/starmap/learning.html>, particularly N. Scott Urquhart, "Troublesome Concepts in Statistics: r² and Power," <http://www.stat.colostate.edu/~nsu/starmap/pps/AP.Stat.Teachers.pps> (accessed approximately March 1, 2011), and the Sequences of Graphs to Illustrate R² that follow. These conditions are derivable from the formula for R², so and can be anticipated and addressed if R² appears artificially high (or low). (c) R² increases as the number of regression variables are added to the model. But this is well known and can be adjusted for. See Roman L. Weil, Peter B. Frank, Christian W. Hughes, Michael J. Wagner, *Litigation Services Handbook: The Role of the Financial Expert* (Hoboken, NJ: John Wiley & Sons, Inc., 2007), 23, <http://books.google.com/books?id=NCCAGTqRjIEC> (accessed approximately March 1, 2011).

⁸¹ A special note of appreciation must be provided to Dr. LaFrancis Rodgers-Rose, who provided insight into the use of the correlation coefficient, in lieu of the more appropriate coefficient of determination, by Educational Testing Service and its reporting of the validity of the SAT many years ago. I was in one of her sociology classes and have been very mindful of the abuse of statistics, particularly the correlation coefficient, ever since. I have no active memory of the evidence she marshaled. Her lecture, however, provided a seed for much of the statistical research completed in this communiqué.

LSAT, $r = 0.3$ translates to an $r^2 (=R^2)$ of 0.09. Hence, instead of the *false* impression that the LSAT is correlated with FYA at 30%, based on r , the casual reader would receive the *truthful* message that there is a 9% association of some sort, if R^2 were published by LSAC. Clearly, this tiny 9% figure would be much harder to defend – or sell. Hence LSAC has an incentive to publish a distortion.

Admissions officers should re-examine their interpretation of LSAC's correlation numbers to determine if their views have been distorted by LSAC.

5.2.1.5. *Gauging standards for correlation and the state of behavioral sciences*

As noted above, there has been ongoing discussion and no consensus about how correlation values reflect strength of correlation. What does seem clear is that researchers are willing to accept lower values of correlation in the social sciences than in the physical sciences. Financial instruments, however, employ a standard closer to the physical sciences. R^2 is a significant factor in determining the quality with which a benchmark index (e.g., S&P 500) predicts a security's or fund's movement. Routinely, $R^2 > 0.85$ (or $r > 0.92$) is considered strongly correlated, while $R^2 < .70$ or ($|r| < 0.83$) is considered weakly correlated.⁸²

While not completely applicable, the econometrics approach must be seen as providing guidance. Human nature is still a significant part of rises and falls on Wall Street. So, it would appear that financial services and econometrics calculations are much more closely aligned with behavioral sciences than with hard sciences.

Statistics pioneer Jacob Cohen considered the question of comparatively low correlations in the social sciences and actually concluded that correlations measured at $r > 0.5$ should be regarded as having a “large effect size” in social sciences and $r = 0.3$ as having a “medium effect size.”⁸³ His reason, however, was that these values characterize the typical ranges of what the modern social sciences find, while physical sciences tend to see $r > 0.99$.⁸⁴ He concluded in 1982:

The fact is that the state of development of much of behavioral science is such that not very much variance in the dependent variable can be predicted. This is essentially merely another way to state the obvious: that behavioral sciences collectively are not as far advanced as the physical sciences.⁸⁵

This statement, which would seem to be axiomatic, is actually a stunning indictment of the quality of research in behavioral sciences. Cohen did not term these correlation thresholds “good” or adequate measures of fit. He directly acknowledged that these values are low with “not very much variance in the dependent variable can be predicted.”

⁸² Investopedia Contributors, “R-Squared,” *Investopedia*, <http://www.investopedia.com/terms/r/r-squared.asp> (accessed approximately March 1, 2011).

⁸³ Cohen, *Statistical power analysis*, 78-81.

⁸⁴ Cohen, *Statistical power analysis*, 78.

⁸⁵ Cohen, *Statistical power analysis*, 78.

We are forced to ask, however, whether models where “not very much variance in the dependent variable can be predicted” are acceptable as a significant part of the law school admissions decision and whether the LSAT behavioral science should have made more progress in the last 30 years.

Some statisticians are more direct, clearly stating that a low r or R^2 indicates that there is minimal correlation, even in the social sciences. When citing an example SAT correlation of 0.40 to introduce the definition of coefficient of alienation, authors Kaplan and Saccuzzo state⁸⁶:

For example, if the correlation between the SAT score and performance in the first year of college is .40, the coefficient of determination is .16. ... For the SAT example, the coefficient of alienation is $\sqrt{1 - .16} = \sqrt{.84} = .92$. **This means that there is a high degree of nonassociation between SAT scores and college performance.** (emphasis added)

5.2.1.6. *LSAT’s historical data revisited, using coefficients of determination and alienation*

To better address the key questions indicated above, we capture the correlation coefficients from several LSAC studies for as much of the last 30 years as is available online and calculate both R^2 and the coefficient of alienation.

⁸⁶ Kaplan and Saccuzzo, *Psychological Testing*, 85.

5.2.1.6.1. How much of FYA is NOT explained by LSAT – National Correlation Studies

Table 7: How Much of FYA Is Not Explained by LSAT – Correlation Studies from 1995 – 2008

Based on *Predictive Validity of the LSAT from 1995 - 2008*

LSAC Studies: National Summaries of Correlation Studies	Calculations based on r		Calculations based on R^2	
	LSAC Published Value of Correlation Coefficient (r)	% of Variation in FYA NOT Explained by Model: Coefficient of Alienation $\sqrt{1-r^2}$	% of Variance in FYA Explained by Model: Coefficient of Determination (R^2)	% of Variance in FYA NOT Explained by Model: Coefficient of Determination ($1 - R^2$)
1995	0.40	91.65%	16.00%	84.00%
1996	0.40	91.65%	16.00%	84.00%
1997	0.39	92.08%	15.21%	84.79%
1998	0.39	92.08%	15.21%	84.79%
1999	0.39	92.08%	15.21%	84.79%
2000	0.40	91.65%	16.00%	84.00%
2001	0.40	91.65%	16.00%	84.00%
2002	0.39	92.08%	15.21%	84.79%
2003	0.37	92.90%	13.69%	86.31%
2004	0.35	93.67%	12.25%	87.75%
2005	0.34	94.04%	11.56%	88.44%
2006	0.33	94.40%	10.89%	89.11%
2007	0.32	94.74%	10.24%	89.76%
2008	0.33	94.40%	10.89%	89.11%

Notes:

1. R^2 explains the % of variation in FYA attributable to LSAT. $(1 - R^2)$ is the % of variation in FYA unexplained by the LSAT. The power of 2 is native to these estimates (which are on the order of variance and sum of squares calculations).
2. The Coefficient of Alienation (COA) is a complement to r , on the order of r . Together, in theory, they explain all of the variation in FYA. However, COA is little more than a poor, linear approximation of $(1 - R^2)$, so contains erroneous information. COA is provided, nonetheless, in deference to the LSAC and its choice to provide correlation coefficient (r), instead of R^2 . The COA indicates even *less* of an association between LSAT score and FYA than does R^2 .

Index:

Correlation coefficients were taken from the following studies, which are part of the LSAC series *National Summaries of Correlation Studies*, updated biannually:

1. Predictive Validity of the LSAT: A National Summary of the 1995–1996 Correlation Studies (TR 97-01)

2. Predictive Validity of the LSAT: A National Summary of the 1997–1998 Correlation Studies (TR 99-01)
3. Predictive Validity of the LSAT: A National Summary of the 1999–2000 Correlation Studies (TR 01-02)
4. Predictive Validity of the LSAT: A National Summary of the 2001–2002 Correlation Studies (TR 03-01)
5. Predictive Validity of the LSAT: A National Summary of the 2003–2004 Correlation Studies (TR 05-02)
6. Predictive Validity of the LSAT: A National Summary of the 2005–2006 Correlation Studies (TR 07-02)
7. Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies (TR 09-03)

5.2.1.6.2. How much of FYA is NOT explained by LSAT – Differential Prediction for Racial/Ethnic Subgroups

Table 8: How Much of FYA Is Not Explained by LSAT – Race/Ethnicity 2005 - 2007
Based on *Differential Prediction for Racial/Ethnic Subgroups from 2005 - 2007*

LSAC Studies: National Summaries of Correlation Studies	Calculations based on r		Calculations based on R ²	
	LSAC Published Value of Correlation Coefficient (r)	% of Variation in FYA NOT Explained by Model: Coefficient of Alienation $\sqrt{1-r^2}$	% of Variance in FYA Explained by Model: Coefficient of Determination (R ²)	% of Variance in FYA NOT Explained by Model: Coefficient of Determination (1 - R ²)
2005 – 7				
Asian American	0.22	97.55%	4.84%	95.16%
Black	0.26	96.56%	6.76%	93.24%
Latino	0.32	94.74%	10.24%	89.76%
White	0.23	97.32%	5.29%	94.71%

Notes: (same as above)

Index:

Correlation coefficients were taken from the following studies, which are part of the LSAC series *National Summaries of Correlation Studies*:

- *Analysis of Differential Prediction of Law School Performance by Racial/Ethnic Subgroups Based on 2005–2007 Entering Law School Classes* (TR 09-02)

5.2.1.6.3. How much of FYA is NOT explained by LSAT – Differential Prediction for Gender Subgroups

Table 9: How Much of FYA Is Not Explained by LSAT – Gender 2005 - 2007
Based on *Differential Prediction for Gender Subgroups from 2005 – 2007*

LSAC Studies: National Summaries of Correlation Studies	Calculations based on r		Calculations based on R ²	
	LSAC Published Value of Correlation Coefficient (r)	% of Variation in FYA NOT Explained by Model: Coefficient of Alienation $\sqrt{1-r^2}$	% of Variance in FYA Explained by Model: Coefficient of Determination (R ²)	% of Variance in FYA NOT Explained by Model: Coefficient of Determination (1 - R ²)
2005 - 7				
Male	0.29	95.70%	8.41%	91.59%
Female	0.38	92.50%	14.44%	85.56%

Notes: (same as above)

Index:

Correlation coefficients were taken from the following studies, which are part of the LSAC series *Differential Prediction for Gender Subgroups*, updated biannually (the correlation coefficients were published in only one study):

- *Analysis of Differential Prediction of Law School Performance by Gender Subgroups Based on 2005–2007 Entering Law School Classes* (TR 10-01)

5.2.1.6.4. How much of FYA is NOT explained by LSAT – Item Type Validity

Table 10: How Much of FYA Is Not Explained by LSAT – Item Type 1993 - 1996
Based on *Item Type Validity*
13 tests from June 1993 – June 1996

LSAC Studies: National Summaries of Correlation Studies	Calculations based on r		Calculations based on R ²	
	LSAC Published Value of Correlation Coefficient (r)	% of Variation in FYA NOT Explained by Model: Coefficient of Alienation $\sqrt{1-r^2}$	% of Variance in FYA Explained by Model: Coefficient of Determination (R ²)	% of Variance in FYA NOT Explained by Model: Coefficient of Determination (1 - R ²)
1993 – 6				
Analytic Reasoning	.244	96.98%	5.95%	94.05%
Logical Reasoning	.358	93.37%	12.82%	87.18%
Reading Comp	.314	94.94%	9.86%	90.14%
Total Raw Score	.399	91.70%	15.92%	84.08%
LSAT Scaled Score	.388	92.17%	15.05%	84.95%

Notes: (same as above)

1. Same notes as above
2. AR – “Analytic Reasoning”
3. LR – “Logical Reasoning”
4. RC – “Reading Comp”

Index:

Correlation coefficients were taken from the following studies:

- *LSAT Item-Type Validity Study (TR 98-01)*

5.2.1.6.5. Observations about LSAT’s historical performance as a measure of FYA

An examination of these tables shows several disturbing trends:

1. LSAT correlations with FYA have been dropping significantly since 2000, down from 0.40 to 0.33. This corresponds to a drop in the percentage of the variation in FYA that is **NOT** explained by the LSAT from 84% to 89%.
2. Among males, LSAT correlation with FYA is only 0.29. This means that the LSAT explains only 8.41% of the variation in FYA scores and leaves 91.59% of the variation in FYA unexplained.

3. The Analytic Reasoning section's correlation with FYA is 0.244. This means that the Analytical Reasoning section explains only 5.95% of the variation in FYA scores and leaves 94.05% of the variation in FYA unexplained.
4. Among Asian Americans, LSAT correlation with FYA is only 0.22. This means that the LSAT explains only 4.84% of the variation in FYA scores and leaves 95.16% of the variation in FYA unexplained.
5. Among Whites, LSAT correlation with FYA is only 0.23. This means that the LSAT explains only 5.29% of the variation in FYA scores and leaves 94.71% of the variation in FYA unexplained.
6. There is inconsistency in correlations calculated by different researchers attempting to address different aspects of the LSAT.

5.2.1.6.6. *The psychology vs. responsibility of declaring LSAT's results poor and unacceptable*

If we accept Cohen's rule-of-thumb thresholds, the LSAT correlations demonstrate that the LSAT has a small to medium effect on FYA. If we look at these from an econometrics standpoint, the LSAT has negligible correlation with FYA and would never trigger a "buy" decision. Should we trust law school admissions decisions to any less precision than we do our investments? If so, how much less precise is acceptable – 5%, 25%, 50%? Note that the correlations above are 80% less accurate than investment thresholds.

Many universities, the ABA, statisticians and researchers will insist that r values between 0.2 and 0.5 are acceptable measures, despite the evidence to the contrary (corresponding R^2 values are 4% - 25%). But, their defense of these thresholds could have much more to do with a pattern-driven illusion described by Cohen. He stated⁸⁷:

Thus, when we consider $r = .50$ a large Effect Size, the implication that .25 of the variance accounted for is a large proportion of what must be understood *relatively*, not absolutely. ... The question, "relative to what?" is not answerable concretely. The frame of reference is the writer's subjective averaging of PVs [proportions of variance] from his reading of the research literature in behavioral science. Since no one reads a stratified random probability sample of the behavioral science literature (whose definition alone would be no mean task), this average may be biased in a "soft" direction, i.e., towards personality-social psychology, sociology, and cultural anthropology and away from experimental and physiological psychology.

Those who have lived with the issues surrounding the LSAT for many years should re-examine their support, with this in mind.

Accreditation bodies, universities, researchers and policymakers have a duty to confront directly the incontrovertible truth that the LSAT fails to explain 85% - 95% of what it is designed, required and relied upon to do. The state of educational testing has clearly not improved over the 30 years since Cohen analyzed it. Correspondingly, we do *not* have to

⁸⁷ Cohen, *Statistical power analysis*, 78-9.

continue to penalize a fresh crop of prospective lawyers every year by subjecting them to the arduous preparation and then judgment of a test that simply fails to have any meaningful statistical merit. In fact, the precipitous decline in LSAT correlations since 2000 suggests that matters are becoming worse. The errors arising from admissions decisions which are based, in part, upon the LSAT are only growing.

It is possible that the LSAT is measuring something other than innate reasoning or reading comprehension ability. Throughout this document, we suggest that the LSAT is measuring, to a great extent, the capacity, in time and resources, of test-takers to prepare for it. But, irrespective of what it measures, it is clear that LSAC's own research indicates that it is a failure. Perhaps the only thing that keeps it afloat is the deceptive use of correlation coefficient and the continued acceptance by interested parties that poor levels of correlation are somehow the best that we can do. Given the distortions that low correlations introduce, it seems clear, however, that the LSAT does much more harm than good. This is even more emphatically demonstrated when one considers the anxiety this test causes tens of thousands of prospective law school students every year, as well as the inequities that preparation for this test introduces.

We do not have to rely on patterns of poor results in educational testing as the basis for our interpretations. We are required, and the state of educational testing demands, that we use our own common sense, life experience, knowledge of statistics, sense of fairness and whatever else makes us human to reach our own conclusions about the LSAT and the harm that it does to prospective lawyers, their families, law schools and the legal system.

5.2.2. LSAC researchers repeatedly draw unjustified conclusions using correlation coefficient (r)

In every LSAC study, the researchers falsely attribute an acceptable predictability level to low r values. It appears to be a spin requirement of sorts. Below, we examine some of these claims.

5.2.2.1. *Predictive Validity of the LSAT*

In *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*, at <http://www.lsac.org/LsacResources/Research/TR/tech-reports.asp>, the researchers stated that:

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. (p. 8)

However, the correlation coefficients of **undergraduate GPA predictability of FYA** and **LSAT predictability of FYA** differed by only 0.04 in 2007 and 0.05 in 2008. The corresponding mean R^2 differences were only 2.4% $= (0.32^2 - 0.28^2)$ in 2007 and only 3.1% $= (0.33^2 - 0.28^2)$ in 2008. That the LSAT explains 3% more variation in FYA than undergraduate GPA is negligible, particularly when one considers that there is no evidence that this difference is concentrated in any narrow portion of the LSAT, undergraduate or FYA value sets.

Even worse is that the researchers would dare to compare measures that are, in and of themselves, poorly correlated with FYA. The mean undergraduate GPA correlation is 0.28 and LSAT correlations are between 0.32 and 0.33, for the years in question. The corresponding R^2 values are 7.8% and 10.2%/10.9%. All of these R^2 values fall well below any conceptual notion of a strong association or correlation with FYA. Comparisons between them are dubious, at best.

5.2.2.2. *Analysis of Differential Prediction of Law School Performance - Racial/Ethnic Subgroups*

Similarly, in *Analysis of Differential Prediction of Law School Performance by Racial/Ethnic Subgroups Based on 2005–2007 Entering Law School Classes (TR 09-02)*,⁸⁸ correlation coefficients of 0.22, 0.26, 0.32, 0.23 are presented for Asian-Americans, African-Americans, Latinos and Whites, respectively.⁸⁹ Per the above, none of these are strongly correlated with FYA. Combining LSAT and undergraduate GPA produces correlations for these subgroups that range from 0.39 to 0.46, again all weakly correlated.⁹⁰ The mean errors between actual and predicted FYA's at law schools for each group are later compared⁹¹ in an attempt to demonstrate the validity of the LSAT. However, the mean error statistic clearly obscures the variation in range within any of these groups at any law school and across all law schools. And, given the low correlations, there is likely to be significant variation about the predicted values.⁹²

5.2.2.3. *Conclusions on LSAC's interpretation of correlation coefficient*

The LSAC's use of the correlation coefficient, r , is a cheap parlor trick that masks the poor performance of the LSAT in predicting FYA or anything else about law school, bar passage rates, etc. When converted to the direct measure, R^2 , the LSAT-FYA relationship is revealed to be so weakly correlated as to be negligible. Combining LSAT with undergraduate GPA produces slightly higher, but still weakly correlated predictors of FYA. LSAC must be aware of the deception caused by its exploitation of r . But, like an unregulated drug dealer, it will keep supplying this intoxicant until we reject it or reject the LSAC.

⁸⁸ See Lynne L. Norton, Deborah A. Suto, Lynda M. Reese, *Analysis of Differential Prediction of Law School Performance by Racial/Ethnic Subgroups Based on 2005–2007 Entering Law School Classes, LSAT Technical Report 09-02* (Newton, PA: Law School Admission Council, October 2009), <http://www.lsac.org/LsacResources/Research/TR/TR-09-02.pdf> (accessed approximately March 1, 2011).

⁸⁹ Norton, et. al., *Analysis of Differential Prediction 09-02*, 14, Table 6.

⁹⁰ Norton, et. al., *Analysis of Differential Prediction 09-02*, 14.

⁹¹ Norton, et. al., *Analysis of Differential Prediction 09-02*, 15, Table 7.

⁹² The residual calculations are essentially a sum of squares operation, without the squares. Positive and negative residuals cancel, thus giving a lower error value than a true sum of squares would. It is unclear what statistic this becomes, but it is not a measure of the *magnitude* of the errors. That the mean of these diminished, imperfect measures is then taken across law schools produces more uncertainty. The high variability in the models is obscured by faulty residual calculation. Since each of the law school calculations has a different variance, but there is no explicit compensation for heteroscedasticity in the final mean operation, the final calculation is imprecise and unrecognizable, probably useless as a basis for inferences.

5.3. Clear refusal to investigate questions damaging to the LSAT

Any reader of the LSAC studies immediately ponders obvious questions that the researchers never bothered to ask. Virtually all of those questions would shed light on the usefulness of the LSAT, but it would appear that they could also be extremely damaging to the LSAT. LSAC should be forced to address these questions.

5.3.1. Missing analysis of how LSAT preparation impacts LSAT performance

In *Summary of Self-Reported Methods of Test Preparation by LSAT Takers for Testing Years 2005–2006 through 2007–2008*,⁹³ the following questions are *never* pursued – or at least never published:

- How many hours per day did respondents spend studying for the LSAT?
- How many total hours did respondents spend studying for the LSAT?
- How many elapsed calendar days, weeks or months did respondents spend studying for the LSAT?
- What percentage of respondent’s preparation time was spent on each LSAT section?
- What kinds of anxieties, if any, did respondents feel, while preparing for each LSAT section?
- For each LSAT section, how well did respondent’s academic background prepare him/her?
- How much money did respondents spend studying for the LSAT?
- How much time missed at work or school or away from family was required to prepare for the LSAT?
- How much value did respondents get from each type of preparation?
- How was the amount of preparation time correlated with LSAT score?
- How was the amount of money spent for preparation correlated with LSAT score?
- How were the types of preparation methods correlated with LSAT score (individually, multiply – this is a complex question, but those who utilized only one method would reveal clearest results)?
- How much, if any, anxiety did respondents experience at various points in their preparation (when first examining a sample LSAT, mid-way through their preparation, after completing a preparation course, the last few days of their preparation)?
- What types of stressors/factors impacted their preparation and to what extent: finances, care for children, care for adult, personal illness, employment, educational attainment, length of time since completing most recent degree, availability of test preparation course, etc.?
- What did respondents feel on the day of the exam?
- What sacrifices did respondents endure to take the actual exam?
- How did respondents feel immediately after taking the exam?
- How did respondents feel after taking the exam, while waiting for their scores?
- How are many of these factors correlated with LSAT score?⁹⁴

⁹³ See Josiah Evans, Andrea E. Thornton, Lynda M. Reese, *Summary of Self-Reported Methods of Test Preparation by LSAT Takers for Testing Years 2005–2006 through 2007–2008*, LSAT Technical Report 08-04 (Newton, PA: Law School Admission Council, October 2008), <http://www.lsac.org/LsacResources/Research/TR/TR-08-04.pdf> (accessed approximately March 1, 2011).

These are the types of questions which could easily undermine the desire of law schools and government officials to continue to sanction the use of the LSAT in admissions. It is likely that the LSAC knows this. But, they should be pressed to release any such information if they have it or to develop what they lack.

5.3.1.1. *LSAT scores decline with age – reasons never explored by LSAC*

As noted above, there is a disturbing, but telling trend captured in *Analysis of Law School Applicants by Age Group: ABA Applicants 2005-2009*.⁹⁵ The researchers present data that suggests that the mean LSAT scores decline precipitously with age, from a mean of 155 for the 25 and under subgroup to a mean of 144, for the 40+ subgroup.⁹⁶ A natural question arises that was never asked by the researchers. Why?

5.3.1.2. *LSAT scores decline with age – implications for preparedness*

As noted above, one explanation that should have been explored was that as one ages, she is more likely to have family responsibilities and face financial pressures. Therefore, the older prospective student is likely to lack the preparation time and/or money (for study materials and commercial test preparation classes) that the younger test-taker has. Hence, the older student is likely not to be as prepared for the LSAT and would perform more poorly. The LSAC likely did not ask this obvious question because the answer could reveal that performance on the LSAT – in part due to its *speededness* – is related most strongly to how much preparation time and money test-takers have, not ability or drive to perform well in law school.⁹⁷

5.3.2. Use of mean when median and percentiles would be more informative

Another set of inappropriate omissions are median and percentile correlation scores. In all but a small number of studies, LSAC calculates mean correlations and LSAT scores, not medians or percentiles. This deprives the reader of any measure of how the outliers affect the results.

⁹⁴ The mean LSAT score and mean age are given for any respondent reported using a given preparation method (or none). But, this exaggerates the effect of outliers and deprives the reader of the useful information that medians, quartiles and other percentiles would provide. As well, it does not attempt to demonstrate any correlations between a given preparation technique and LSAT score, any cross-correlations among methods, or any multiple-correlations.

⁹⁵ See Dustman and Handwerk, “Analysis of Law School Applicants,” 8.

⁹⁶ There could be other factors that partially explain the data, such as the greater number of African-American test-takers among older subgroups. (African-Americans consistently have a lower performance than most other groups for reasons that the LSAC has never explained.) As well, some of this could be explained by the researchers use of the mean and not median or quartiles (outliers with unusual life-responsibilities could be weighing down the average).

⁹⁷ The test could also be measuring other factors faced by older populations, such as less attentiveness during their allotted preparation time due to family demands and distractions, deteriorating health (fading eyesight, hearing, etc.), nagging concerns about whether law school is even a viable option for them. We do not know which, if any of these factors impact the older student, of course, because the LSAC has not studied them. It is extremely unlikely that they could explain the magnitude of the declines, however, per the research cited in footnote #25.

5.3.2.1. *Misleading analysis of repeat test-takers*

5.3.2.1.1. *Performance study of 2000 – 2007 repeaters*

In *The Performance of Repeat Test Takers On the Law School Admission Test: 2000–2001 Through 2006–2007 Testing Years (TR 08-01)*,⁹⁸ numerous obfuscations are manifested to the benefit of the LSAT’s reputation, but which are misleading and harm the perception of those who score at the lower ranges of the LSAT.

The study asserts that:

Across testing years, the mean LSAT scores were highest for first-time test takers (151.1), lower for second-time test takers (150.2), and lowest for third-time test takers (146.8). ... Test takers who repeated the LSAT gained an average of 2.8 points the second time they took the test and 1.9 points the third time they took the test (compared to the second time). Despite these mean score gains, their mean LSAT scores are still usually lower than the mean for first-time test takers, as indicated above and in more detail in the report. (p. 1)

This gives the impression that repeating the LSAT has little effect on the score and suggests that the LSAT is measuring some intrinsic quality of the test-taker. In fact, the use of the mean simply mixes the performance of those incapable of much additional preparation since their earlier exam(s) with those who actually were able to spend much more time preparing, those who learned more test-taking skills, those who simply learned to exploit a weakness in the test that many of the one-time test takers did, The use of the mean makes it impossible to tell whether re-taking the exam is universally fruitless or sometimes extremely beneficial. Stating that there was an improvement of only 1.9 - 2.8 points would seem to suggest that the LSAT is consistently an outstanding measure of something intrinsic about the test-taker and re-taking the exam will only reveal the same. Of course, this could be false.

Unfortunately, nowhere in the study are raw scores, medians or percentiles presented, so the illusion created by the LSAC is complete. Stating the results of the medians and percentiles would have painted a far more complete picture – this was particularly important to depict in the Executive Summary (from which the above citation was extracted).

This pattern is present in the other two studies of the same type on the LSAC Research page.⁹⁹

⁹⁸ See Thornton, *et. al.*, “The Performance of Repeat Test Takers 08-01.”

⁹⁹ See Jennifer R. Duffy, Susan P. Dalessandro, Lisa Anthony Stilwell and Kimberly A. Swygert, *The Performance of Repeat Test Takers On the Law School Admission Test: 1994–1995 Through 2000–2001 Testing Years, LSAT Technical Report 01-03* (Newton, PA: Law School Admission Council, 2001), <http://www.lsac.org/LSACResources/Research/TR/TR-01-03.asp> (accessed approximately March 1, 2011) and Andrea E. Thornton, Deborah A. Suto, Jennifer R. Duffy, and Lynda M. Reese, *The Performance of Repeat Test Takers On the Law School Admission Test: 1997–1998 Through 2003–2004 Testing Years, LSAT Technical Report 04-02* (Newton, PA: Law School Admission Council, 2004), <http://www.lsac.org/LSACResources/Research/TR/TR-04-02.asp> (accessed approximately March 1, 2011).

Because data for racial and gender subgroups are provided, reporting only the mean for each group can give the impression that all members of one group perform more poorly than all members of another. While this could have been avoided with raw, median and percentile numbers for each sub-group, LSAC chose to provide what can only be seen as an inadequate caveat at the very end of the Executive Summary:

Also note that summary statistics across gender or race/ethnicity describe subgroup differences, not individual differences. Thus, for example, a repeat test taker from one racial/ethnic group may outperform 90% of the repeat test takers from another racial/ethnic group, even though the subgroup mean differences might suggest otherwise. (p. 1)

5.3.2.1.2. *Information hidden in RepeaterData.pdf*

As noted previously, LSAC maintains a data file, called **RepeaterData.pdf**,¹⁰⁰ which tracks test scores of those who repeat the LSAT. It lists ranges of the latest scores, in increments of 10, received by those who repeated the LSAT in the most recent test administration year. Unfortunately, because the range of new test scores is fixed at 10 scaled points, granular analysis is not possible. However, by closely examining the range data at regular intervals, we were able to discern important facts about 2 of 10 non-overlapping sections of the data:

- Section A: Approximately 39% of repeat test-takers improved their most recent scores by 4 – 13 scaled points (this corresponds to 4 – 23 additional correct answers);
- Section B: Approximately 13% of repeat test-takers improved their most recent scores by 8 – 17 scaled points (corresponding to 8 – 29 additional correct answers, and not counting those in the category above);
- Within the sample of 8,424 repeat test-takers we examined (in Sections A and B), 2,171 (26%) demonstrated improvement in one of the two score ranges indicated above.

Since Sections A and B were chosen at random, this is likely representative of the entire set of repeat data-takers. The results refute LSAC's claim that repeat testers only marginally improve their test scores. In fact, for a significant number of test-takers, the improvements are dramatic. Recovering from illness or bad days are unlikely to be the reasons for the improvement in so many people from throughout the test administration regions. Additional preparation is the likely reason for the improvement. That additional preparation can make a prospective student, who initially appeared barely qualified to attend law school, suddenly look like an early admission candidate to one of the top law schools suggests that the LSAT does not consistently measure ability.

Clearly, such transformations, often over the course of a few weeks or months, would not add to the credibility of the LSAT as a stable predictor of law school performance and would suggest that the LSAT is really just measuring how much preparation time and money prospective students have. LSAC should be encouraged to publish complete data, by percentile, for repeat test-takers.

¹⁰⁰ LSAC, "LSAT Repeater Data."

5.3.2.1.3. *Validity of Law School Admission Test Scores for Repeat Test Takers*
In *The Validity of Law School Admission Test Scores for Repeat Test Takers: 2005 Through 2008 Entering Law School Classes (TR 10-02)*,¹⁰¹ and frequently offered as some measure of the stability or validity of LSAT scores, the researchers conclude:

Despite the fact that repeat test takers, on average, increase their LSAT performance, even their increased scores tend to be lower than those earned by onetime test takers at their law school. ... The magnitude of the difference varies among individual schools, but only 19 schools showed a mean for repeat test takers that exceeded the one-time test takers' mean score (p. 12)

This is extremely misleading. Those who repeat the exam were clearly dissatisfied with their previous test result(s), whether through lack of preparation, hardship, illness, chance, etc. Naturally, their initial test scores would be lower than those who were satisfied. Additional test preparation should help to improve their scores, but there is no reason to believe that the underlying structural issues of their lives which prevented adequate preparation for their first LSAT would *ever* fully evaporate. So, there is every reason to believe that they might never “catch up to” the one-time test takers.

The study should have looked at the underlying factors, such as preparedness, controlled for them in both the one-time and repeat test taker groups, then drawn conclusions. Because none of this was done, this study could have been simply measuring the persistence of lack of test preparation time in a population of test-takers eager to do well, despite hardship. But, we are not provided such information that would redeem the test takers and, instead, are fed information that, superficially, redeems the test.

The use of the mean is also misleading, as indicated above, because it exaggerates the effect of outliers, but masks their existence. We don't know how many repeat test-takers actually do far better than the one-time test-takers or whether repeating the test dramatically improved the results for a significant number of prospective students. Such information could show whether LSAT scores are really dictated by how much time and money (for test materials and commercial classes) a test-taker spends, as opposed to ability or work ethic upon entering law school. Use of the median, with percentile breakdowns, would have helped to answer some of these questions.

Additionally, the study appears to show that using the highest test score, lowest, most recent or average test score for repeat test-takers results in over-predictions and under-predictions of FYA that are far more pronounced than for the single test-takers. However, the difference in over-under prediction is almost exclusively between 1 and 2 grade points on a scale of 50, or just 2 – 4%. This would appear to be negligible error. And, if such differences do matter to law schools, then it should also be posited that what

¹⁰¹ See Andrea Thornton Sweeney, Laura A. Marcus, Lynda M. Reese, *The Validity of Law School Admission Test Scores for Repeat Test Takers: 2005 Through 2008 Entering Law School Classes, LSAT Technical Report 10-02* (Newton, PA: Law School Admission Council, October 2010), <http://www.lsac.org/LSACResources/Research/TR/TR-10-02.asp> (accessed approximately March 1, 2011).

the study really measures is the extent to which the barriers to LSAT preparedness persist and negatively impact performance in law school.

This is probably not the story that the LSAC wants to tell, however, because it suggests that the LSAT itself is just a measure of who has the time and money to prepare.

5.3.2.1.4. *Dramatic differences in average LSAT scores raise questions about utility of LSAT*

Perhaps the greatest omission from the study *The Validity of Law School Admission Test Scores for Repeat Test Takers: 2005 Through 2008 Entering Law School Classes (TR 10-02)*,¹⁰² is that a stunning result of the study is not highlighted in a significant way. The study notes, in passing:

The distribution presented in Figure 1 displays the percentage of repeat test takers at each of the 184 law schools included in this study. Repeat test takers comprised 14–59% of the entering classes studied. ... **The range of mean LSAT scores across schools for the combined group of one-time and repeat test takers is fairly substantial, varying from a low mean of 135.4 to a high mean of 170.3.** (p. 6, emphasis added)

Similarly, from *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*¹⁰³, the range of LSAT means among universities for *all students*, not just repeat test-takers is from 135.33 to 170.34, in 2008.

This has to be considered earth-shattering. The fact that all of these law schools are graduating competent lawyers (certainly there have been no media accounts to the contrary), yet exhibit a median range of LSAT scores from 135 to 170 for repeat test-takers and their general population alike, suggests that a high LSAT score currently has little to do with matriculation at many schools and, perhaps, has little to do with graduation, bar passage rates, obtaining employment after law school, or success as an attorney. After appropriate analysis, these could be yet additional reasons to abolish the LSAT.

That there is a dramatic, 35-point difference in LSAT scores among law schools is troubling from an accreditation perspective as well. How can an accrediting agency mandate LSAT use for all law schools when the experience of each school has been shown to be dramatically different? Does this one-size-fits-all policy help the schools with a 135 point mean select students? Surely, for those schools, the LSAT has no practical value. One could simply assume that their average student had neither the time nor resources to do any meaningful LSAT preparation, so the LSAT score is totally useless. Why, then, should the LSAT be mandated for such students and schools by any accrediting agency? Is that not cruel and unusual punishment for the students? It is not just an invitation to the imposition of unnecessary and unflattering characterizations of the schools?

¹⁰² See Sweeney, *The Validity of Law School Admission Test 10-02*.

¹⁰³ See Stilwell, *Predictive Validity of the LSAT 09-03*.

6. PROPOSAL C. LSAT'S PREDICTABILITY RATES OF FYA JUSTIFY RESCINDING REQUIREMENTS THAT STUDENTS TAKE THE LSAT AND RECOMMEND A 100% COMPENSATORY MODEL.

6.1. LSAT's poor performance cannot justify accreditation mandate for LSAT

As indicated above, the LSAC's latest national study of the predictability of the LSAT, entitled *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*, concludes that, at the average law school, the LSAT correlation with FYA is 0.33 (as measured by the correlation r , with a standard deviation of 0.10).¹⁰⁴ The coefficient of determination, r^2 , is 0.11. This means that the LSAT only explains 11% of the variability in FYA at a typical law school. Note that a flip of a coin will correctly choose among two possible outcomes 50% of the time. (This would be a good comparison only when the LSAT is asked to predict among two grades, pass or fail, etc.)

The study also attempts to apply transformative measures to the data¹⁰⁵, but those calculations contain errors and distortions, as demonstrated in section 10.4.

Fundamentally, by requiring students to take the LSAT, accrediting agencies force prospective students into traumatic experiences for a measure that only explains 11% of the variability in FYA, on average. Would any cost-benefit analysis justify this?

The above study does reveal that some law schools have a higher correlation coefficient, but even the highest shown, $r = 0.56$ ¹⁰⁶, have a coefficient of determination, $r^2 = 0.31$. Those law schools are rejecting students using LSAT scores that correctly predict variation in FYA only 31% of the time. The LSAT being roughly correct only 1/3 of the time would have to be considered unsavory, intuitively, by the general population and the great majority of law school applicants.

Because the quality predictive capability of the LSAT is the rare exception among law schools, if it ever occurs, there is no reason that any accreditation body should mandate that any school require students to take that test.

6.2. Precautions about withdrawing Standard 503 and altering Standard 501

It must also be stated that leaving the decision of whether to require applicants to take the LSAT up to individual law schools, as does the November 2010 proposed revisions to the ABA Standards, only fractures and distributes the problem into 200 smaller pieces. If the ABA and other entities determine that the research really does characterize the LSAT as invalid, then no law school should be permitted to require applicants to take it. We have suggested language above that accomplished this.

¹⁰⁴ Stilwell, *Predictive Validity of the LSAT 09-03*, 8, Table 2.

¹⁰⁵ Stilwell, *Predictive Validity of the LSAT 09-03*, 12, Table 3.

¹⁰⁶ Stilwell, *Predictive Validity of the LSAT 09-03*, 8, Table 2.

6.3. Towards a compensatory model for everyone

A critical paragraph in the above study effectively accepts the failures of the LSAT and suggests that there is an alternative path¹⁰⁷:

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, this model allows a high LSAT score to compensate for a low UGPA or, conversely, a high UGPA to compensate for a low LSAT score when schools are making admission decisions.

This extraordinary statement tells us that the LSAT fails to predict performance for a class of students and that universities know how to evaluate them by looking at “other evidence of their ability to do well in law school.”

The questions, then, are how well those compensatory models predict performance for all students and whether they would be just as effective in the total absence of LSAT scores. What is clear is that LSAT scores can be eliminated for significant portions of prospective students. Accrediting agencies should rescind all requirements that such students endure LSAT preparation and should study whether and how compensatory models could be used in lieu of LSAT scores for *all* students.

One note of concern in today’s compensatory models is that bonuses for race and ethnicity will have to be reduced and, ultimately, eliminated over time, to comply with US Supreme Court requirements, as discussed below. However, diversity could be built into the system through some of the measures in the revised Interpretation 501-1, particularly the “obstacles overcome” criteria. As well, diversity could be helped if high school class rank is factored into the admissions model. Since racial and ethnic segregation is still a reality in the US, these “Texas Top 10% Rule” solutions have been shown to be practical and effective ways of increasing and maintaining diversity.¹⁰⁸ They have also been found constitutional.¹⁰⁹

¹⁰⁷ Stilwell, *Predictive Validity of the LSAT 09-03*, 12.

¹⁰⁸ Wikipedia contributors, "Texas House Bill 588," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Texas_House_Bill_588 (accessed approximately March 1, 2011).

¹⁰⁹ Wikipedia contributors, "Texas House Bill 588."

7. PROPOSAL D. RACIAL AND ETHNIC LSAT DISPARITIES MUST BE ELIMINATED WITHIN THE NEXT TWENTY YEARS OR DIVERSITY EFFORTS MAY BE RULED UNCONSTITUTIONAL.

7.1. Argument

Since 1978, the US Supreme Court has been clear that a university promotes a state interest when looking beyond low standardized test scores and considering factors such as race, ethnicity, gender, geographic home of applicants to create a diverse student population, providing these factors are considered merely plusses in a “narrowly tailored” program. This is directly applicable to the LSAT because the compensatory admission models referenced by LSAC clearly employ such strategies (see discussion of *Grutter*, below).

However, the Court has ruled that there is a time limit on such strategies. Accrediting bodies, law schools and LSAC must begin now to fashion an equitable means of disabling the current LSAT time-bomb or racial and ethnic diversity could become a thing of the past in less than two decades. Virtually every recommendation we have provided in this communiqué applies to all races and ethnicities, but the information developed here can be used just as well to help chart an *imminently* necessary corrective course for the LSAT and how diversity decisions are made in law schools.

7.2. Diversity is compelling, but race/ethnicity LSAT corrections will be unconstitutional

In the landmark *Regents of the University of California v. Bakke*, 438 U.S. 265, 269, 98 S. Ct. 2733, 57 L. Ed. 2d 750 (1978), Justice Powell established that diversity in universities “furthers a compelling state interest.” While striking down quotas, the court held that a:

State has a substantial interest that legitimately may be served by a properly devised admissions program involving the competitive consideration of race and ethnic origin.

In *Grutter v. Bollinger*, 539 U.S. 306, 123 S. Ct. 2325, 156 L. Ed. 2d 304 (2003), the US Supreme Court augmented Justice Powell’s view that diversity on law school campuses was a worthy educational and societal goal¹¹⁰:

The Law School's claim is further bolstered by numerous expert studies and reports showing that such diversity promotes learning outcomes and better prepares students for an increasingly diverse workforce, for society, and for the legal profession. Major American businesses have made clear that the skills needed in today's increasingly global marketplace can only be developed through exposure to widely diverse people, cultures, ideas, and viewpoints. High-ranking retired officers and civilian military leaders assert that a highly qualified, racially diverse officer corps is essential to national security. Moreover, because

¹¹⁰ While Justice Powell was not joined in the diversity portion of his opinion by any other justice, the entire five justice majority signed this portion of the *Grutter* opinion.

universities, and in particular, law schools, represent the training ground for a large number of the Nation's leaders, the path to leadership must be visibly open to talented and qualified individuals of every race and ethnicity. Thus, the Law School has a compelling interest in attaining a diverse student body.

At exactly the same time, the court accepted evidence that race-blind admissions policies today would eviscerate diversity:

... a race-blind admissions system would have a "very dramatic," negative effect on underrepresented minority admissions. App. to Pet. For Cert. 223a. He testified that in 2000, 35 percent of underrepresented minority applicants were admitted. Dr. Raudenbush predicted that if race were not considered, only 10 percent of those applicants would have been admitted. *Ibid.* Under this scenario, underrepresented minority students would have comprised 4 percent of the entering class in 2000 instead of the actual figure of 14.5 percent.

Michigan avoided the virtual elimination of underrepresented minorities by establishing a goal of "assembling a class that is both exceptionally academically qualified and broadly diverse." Hence the court's conclusions that:

With respect to the use of race itself, all underrepresented minority students admitted by the Law School have been deemed qualified.

The use of the LSAT alone did not allow Michigan to discern who among "underrepresented minorities" were "both exceptionally academically qualified." (See excerpts from dissent of Justice Thomas, below for LSAT statistics.) Again, this argues for the use of compensatory models on a more widespread basis.

However, the use of race in a compensatory admission model is on life support and, per Justice O'Connor's opinion for the *Grutter* majority, it has less than 20 years of life remaining:

Accordingly, raceconscious admissions policies must be limited in time. This requirement reflects that racial classifications, however compelling their goals, are potentially so dangerous that they may be employed no more broadly than the interest demands. ... all race-conscious admissions programs have a termination point ... We take the Law School at its word that it would "like nothing better than to find a race-neutral admissions formula" and will terminate its race-conscious admissions program as soon as practicable. ... It has been 25 years since Justice Powell first approved the use of race to further an interest in student body diversity in the context of public higher education. ... We expect that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today. ... In summary, the Equal Protection Clause does not prohibit the Law School's narrowly tailored use of race in admissions decisions to further a compelling interest in obtaining the educational benefits that flow from a diverse student body.

As the next twenty years draws near, we are fast approaching a time when the dissent of Justice Thomas will become law of the land:

In other words, the tests [at Columbia Law School in the early twentieth Century] were adopted with full knowledge of their disparate impact [on Jewish immigrants]. Cf. *DeFunis v. Odegaard*, 416 U. S. 312, 335 (1974) (*per curiam*) (Douglas, J., dissenting).

Similarly no modern law school can claim ignorance of the poor performance of blacks, relatively speaking, on the Law School Admissions Test (LSAT). **Nevertheless, law schools continue to use the test and then attempt to “correct” for black underperformance by using racial discrimination in admissions so as to obtain their aesthetic student body. The Law School’s continued adherence to measures it knows produce racially skewed results is not entitled to deference by this Court.** See Part IV, *supra*. The Law School itself admits that the test is imperfect, as it must, given that it regularly admits students who score at or below 150 (the national median) on the test. (emphasis added)

7.3. Racial/Ethnic LSAT disparities almost impervious to change

The racial disparities in LSAT performance are so stark that it is inconceivable that race and ethnicity will evaporate as differentiating factors over the next twenty, forty or even sixty years. The LSAC study *Comparison of LSAT Performance Among Selected Subgroups (SR 90-01)*, described above, demonstrates this disparity clearly. We can examine trends, however, by looking at LSAC studies from 1993 – 2010.

In *LSAT Performance With Regional, Gender, and Racial/Ethnic Breakdowns: 1993–1994 through 1999–2000 Testing Years (TR 00-01)*,¹¹¹ the following mean LSAT scores were reported for each racial/ethnic group measured for the 1993-4 testing year. (Table 4, p. 13)

¹¹¹ See Susan P. Dalessandro, Lisa Anthony Stilwell and Lynda M. Reese, *LSAT Performance With Regional, Gender, and Racial/Ethnic Breakdowns: 1993–1994 through 1999–2000 Testing Years, LSAT Technical Report 00-01* (Newton, PA: Law School Admission Council, October 2000), 3, <http://www.lsac.org/LSACResources/Research/TR/TR-00-01.asp> (accessed approximately March 1, 2011).

Table 11: Mean LSAT Scores by Race/Ethnicity – 1993 - 1994

Group	LSAT Score
African American	141.96
Native American	148.19
Asian American	150.59
Canadian Aboriginal	150.63
Caucasian	152.86
Hispanic	146.74
Mexican American	147.50
Puerto Rican	138.64
Other	149.58

In *LSAT Performance with Regional, Gender, and Racial/Ethnic Breakdowns: 2003–2004 Through 2009–2010 Testing Years (TR 10-03)*,¹¹² the following mean LSAT scores were reported for each racial/ethnic group measured for the 2009-2010 test year. (Table 4B, p. 20) Note that the subgroup names changed slightly.

Table 12: Mean LSAT Scores by Race/Ethnicity – 2009 - 2010

Group	LSAT Score
American Indian/Alaskan Native	146.89
Asian	152.38
Black/African-American	142.04
Canadian Aboriginal	152.94
Native Hawaiian/Other Pacific Islander	146.42
Hispanic/Latino	146.43
Puerto Rican	138.37
White/Caucasian	152.86
Multiple Ethnicities	150.80

Median and quartile statistics would be far more helpful in gauging the spread of scores among subgroups, but the mean snapshot provided by LSAC should be sufficiently enlightening to illustrate the contentions herein. After a seventeen-year period, only median scores for Asians (and Canadian Aboriginals) have increased (2 points).

However:

- African-Americans are still scoring a full **10 points below** Asians and Whites;
- Puerto Ricans are still scoring more than **14 points lower** than Asians and Whites;
- Hispanics and indigenous populations in the US are scoring nearly **6 points lower** than Whites and Asians.

¹¹² See Dalessandro, *LSAT Performance with Regional 10-03*.

After remaining flat for seventeen years, we cannot expect the LSAT to cease differentiating on the basis of race and ethnicity in the decades to come.

7.4. LSAT modifications or replacements must address preparation inequities and law school objectives

In the above, we recommend several ways to reduce inequities in the LSAT, but these measures could just as well be utilized to eliminate racial, ethnic, age (and perhaps other) disparities.

They are:

- Altering fixed-time LSAT administration so that tests are given on an untimed or extra-time basis. Pearson and the College Board have shown that this can be accomplished on a practical basis and that the testing day need be extended only moderately.
- Recommending that LSAC publish (on its website and, for those without internet access, at cost) full disclosure of everything a prospective student needs to know about what the LSAT tests, how it tests, endless examples, test-taking strategies, all which are aimed at helping test-takers maximize their LSAT scores without spending any money whatsoever.
- Removing the AR and RC sections from the LSAT, they are duplicative, adding no predictive value to the LSAT.
- Introducing power and speed scoring measures in a comprehensive and strategic manner. First, remove the AR and RC sections. Second, give the exams on an extra-time basis (perhaps 30 minutes more each section). Third, expand the number of questions to a number that cannot be reasonably answered by any test-taker. Fourth, report two scores: percent correct of those answered (*Mastery score*) and number answered correctly (*Intensity score*). Slow readers and those suffering preparation inequities would likely do well on the ability measure, while others would do well on the speed score. Law schools desiring diversity could emphasize the ability score and de-emphasize the speed score.
- Abolishing the LSAT altogether and replacing it with a 100% compensatory admission model, where low LSAT scores are largely ignored and high LSAT scores are given only a small bonus.
- Recommending that pre-law programs offer high-quality LSAT prep courses.
- Mandating that law schools that rely upon LSAT scores offer free, high-quality LSAT prep courses to potential applicants.

Additional means, absent modifying or abolishing the LSAT, appear below. Whatever the approach, it must begin without delay.

8. PROPOSAL E. WAYS TO REDUCE LSAT INEQUITIES, OTHER THAN MODIFICATION OR ABOLITION OF THE LSAT.

It is fathomable that altering the LSAT might not be forthcoming in the near future. Should this unpleasant circumstance arise, there are still actions that could be taken by accrediting organizations, law schools and LSAC that are likely to improve the LSAT experience for those suffering the LSAT's inequities.

All interested parties can perform independent analyses of LSAC's claims about the LSAT and about any claim in this communiqué. They can tell the truth about what they find. And, they can use their discoveries to help prospective students better prepare for the LSAT as well as help both test-takers and law schools better understand the test results.

Specifically, three types of studies should be performed and, based on the above evidence, two actions could be taken immediately (subject to the results of their studies).

8.1. Commence independent studies

We have questioned the soundness of LSAC's studies, particularly in light of the reliance on the correlation coefficient, obvious questions they never asked, and their spin control. The following would appear to be reasonable research paths.

First, your own statisticians should review the LSAC studies to gauge their scientific merit, particularly the use of the correlation coefficient (r) as opposed to the more informative coefficient of determination (r^2) and whether r -values as low as 0.3 – 0.5 could be considered “strong” correlations.

Second, you could perform studies (or encourage LSAC to release their own studies, if any) that measure the extent to which the LSAT really measures a prospective student's time and resource availability/commitment instead of a capacity for law school success.

Finally, you should investigate why undergraduate GPA's, according to LSAC studies, do not predict law school performance. If law schools require skills not taught at the undergraduate level, then identify how the undergraduate curriculum could be supplemented or otherwise modified. If, however, law schools employ teaching methods that do not translate into good legal practice, then also propose changes in the law school curriculum.

Essentially, independent study and encouraging LSAC to release all relevant unpublished research will help to demystify the LSAT, measure its real value in the admissions process, and address concerns that LSAC simply ignores or dismisses.

8.2. Immediate actions

There are two steps that, based on what we already know about the LSAT and LSAC's studies, could improve the LSAT experience and how law schools use the LSAT. Additional study might be necessary first, but only minimally.

8.2.1. Create free, high-quality LSAT preparation courses at university level

First, accrediting bodies should encourage pre-law programs to offer high-quality, full credit LSAT preparation courses and mandate that law schools which rely upon the LSAT for admissions decisions offer free, high-quality courses. Perhaps grades from the law school courses could be substituted for the LSAT scores or a modified LSAT could be the final.

The undergraduate courses would eliminate cost and time from the equation of LSAT preparation because pre-law students would simply prepare by taking another one of many full credit courses. Law schools offering high quality courses free of charge eliminates the cost of LSAT preparation for undergraduate test-takers who are not pre-law majors and for test-takers long removed from their school days. They must still find or make the time to attend the free courses, however, so this will not help every prospect. Law schools would benefit because they can become more familiar with the very requirements that they have demanded of their applicants for decades. This might encourage them to become more sensitive to applicant concerns about the test.

8.2.2. Make truthful, continuous and pervasive public service announcements about LSAT

Second, any stigma associated with a low test-score, ignorance about how much LSAT preparation is required, and confusion about what the LSAT is likely measuring could be significantly alleviated through a bold and timely public education/remediation campaign. Interested parties could setup a web site (identified below as http://www.***.org) to gather the anecdotal evidence LSAC refuses to collect and to host its independent studies. Interested parties should then announce clearly and *regularly* a message similar to the following (expanded from Howard University's statement):

The LSAT appears to be most useful in measuring your ability to prepare for unfamiliar challenges in a short span of time, not your aptitude for legal work, your academic background or your intelligence. You will generally do poorly if your life circumstances do not permit you the time to study and the financial means to purchase preparation books, old tests, or commercial courses. Most law schools will attempt to compensate for a low LSAT score by considering other factors, but they grant few such compensatory admissions.

If you believe this policy is unfair and would like to change it, please go to http://www.***.org to register your concerns, your LSAT experience, and/or your ideas for improving the system. LSAT statistical studies can be found at <http://www.lsac.org/LSACResources>. Independent evaluations of those studies can be found at http://www.***.org.

This brutal honesty would declare finally that the LSAT is not an aptitude test of any sort that a typical applicant would recognize. It would better inform those lacking quality preparation opportunities so that they understand precisely what to expect of the preparation process and how to digest their scores. Finally, it would serve to pressure

law schools to view the LSAT more realistically. By having to face the truth about what they are demanding of applicants and how the test is likely to be of dubious value, law schools would likely de-emphasize the LSAT in the admissions process. This set of outcomes would remove many of the inequities associated with the LAST.

9. CONCLUSIONS

We have provided an extensive analysis of the LSAT, its invalidity and its inequities. We have analyzed the November 2010 proposals for modifications to the ABA Standards. We have offered numerous recommendations for how to remedy the shortcomings of each.

LSAC's own research demonstrates that parts of the LSAT are duplicative, that the overall scores are poor indicators of law school performance, and that these indicators have become far less reliable over the past ten years. The LSAC has engaged in a pattern of deception to hide these facts in plain sight. Their studies stop tantalizingly short of uncovering what the LSAT really measures, simply because LSAC has refused to ask obvious question about test preparation strategies, hardships and outcomes.

LSAC states that the purpose of the LSAT is "to measure skills considered essential for success in law school." But, there is evidence that the time limits on the tests are too low, turning the LSAT into a test of speed, not ability. It is also peculiar that what purports to be a consistent measure of ability requires, per conventional wisdom, approximately three months of study to prepare for it. What the LSAT really measures is unclear. What is certain is that those with significant LSAT test-taker's preparation burdens (time constraints, financial challenges, and/or intense personal obligations) suffer diminished opportunities to prepare and cannot demonstrate their maximal ability on the LSAT.

While LSAC will not study this directly, there is data that points decidedly to this thesis. One study reveals that LSAT scores tend to decline with age, just as one would expect their barriers to test preparation to rise. Approximately 25% of repeat test takers obtain significant performance increases, as would happen with increased preparation. One-third of all test-takers utilize three or more methods of test preparation; one-third pay for expensive commercial test preparation courses; 40% buy a comparatively expensive book not published by the LSAT. Anecdotal evidence on blogs and among test tutors attests to the commitments of time and money required to excel on the LSAT. Unfortunately, entire classes of people lack that time, money and/or freedom from personal responsibilities to prepare adequately. Underrepresented minorities, the poor of all races, single parents, older applicants, women are all likely to have characteristics correlated with a higher LSAT test-taker's preparation burden than younger, childless and/or middle income test-takers. *All of these groups also have lower than average LSAT scores.*

Of course, LSAC could measure some of these associations quickly, through fee waiver examinations and surveys, and release the information that they already know. But, the common-sense arguments, anecdotal evidence and the available research should, at minimum, raise red flags among accrediting bodies, law schools, and the entire legal community. Does the LSAT measure ability, speed, both? How do type, duration and quality of preparation affect LSAT scores? Should one's preparation strategy affect a consistent measure of his ability? How much do income and structural life challenges affect one's ability to prepare for the LSAT? Are race, ethnicity, age, income, gender, personal responsibilities, business responsibilities correlated with ability to prepare for the LSAT and with LSAT score?

While more study would be helpful, there is enough already known to justify eliminating the LSAT or at least the Analytical Reasoning (AR) and Reading Comprehension (RC) sections from the admissions process. An untimed or extra-time administration of the LSAT would also enable it to measure ability more than speed. The LSAC notes that most law schools already utilize what are termed “compensatory models” to gauge likely success in law schools in ways that the LSAT fails to measure. Unfortunately, those models are utilized only in exceptional circumstances. However, the varied and troubling issues with the LSAT suggest that these models should be used as the rule, though their use of race- and ethnicity-based advantages must be reduced and, ultimately, eliminated, per *Grutter*.

Finally, because compensatory models are used as the exception in an environment where the LSAT seems to be the dominant player in assessing law school readiness, those admitted under compensatory models might suffer the unnecessary and false stigma that they needed help to be admitted. The truth may very well be that their competitors needed and obtained the real advantages. LSAC, accrediting bodies and law schools must be much more vocal about what the LSAT measures, what it does not, how high scores have been obtained, and how everyone can obtain the best possible score for themselves. No interested party can afford to be passive or silent.

We have demonstrated that:

1. The November 2010 proposed ABA standards do nothing to address the significant issues with today’s LSAT. However, we proposed additional modifications to them that would be of tremendous benefit to the admissions process of every law school in the nation and to every law school applicant.
2. The predictive power of the LSAT’s overall score has been camouflaged in mathematical sleight of hand and wordsmithing. In fact, LSAC’s own data reveals extremely limited, if any, correlation and that chance would frequently be a better predictor than the LSAT.
3. The injury caused by the AR/RC sections and the LSAT as a whole cannot be justified by the alleged benefit of what is claimed to be an objective admissions practice, but which some evidence suggests is not. In fact, continued use of the AR/RC sections, in light of the fact that they are superfluous, is unjust and does not ensure that the most qualified or diverse class enters a given law school.
4. There is considerable variability in LSAT correlations and median LSAT scores among law schools. An accrediting agency cannot universally mandate LSAT use when the experience of each school has been shown to be dramatically different.
5. There is some evidence that the quality of LSAT preparation has a significant effect on LSAT score. This suggests that law school admissions are being impacted by the arbitrariness and inequities associated with the luck of the draw: who has the time and resources to prepare adequately and who does not.
6. Inequities in the test-taker’s preparation process can be partially addressed by going to an untimed or extra-time test administration. These have proven to be practical and result in fairer tests of ability. In an extra-time administration, provide both

power and speed scores; increase the number of questions to ensure that the speed score is relevant and differentiated from the power score.

7. Law schools have adopted compensatory models that predict law school success where the LSAT fails. These models present an alternative to the LSAT's use. While more study would be helpful, there is enough evidence to suggest that accrediting bodies should withdraw their LSAT mandates in favor of a fundamental reliance on compensatory models that are consistent with case law realities. In this scenario, however, the LSAT must be prohibited from being required anywhere. The decision cannot be left to individual schools, because the national failures of the LSAT would simply become 200 local failures and the inequities and injustices would continue unabated.
8. Additional research and the full disclosure of what the LSAT measures, as well as how to ensure that the LSAT score reflects one's abilities must be undertaken by accrediting bodies, LSAC and law schools.

Your own statisticians could verify everything stated in this communiqué by studying the same reports and asking the LSAC tough questions. The LSAC will not tell me how to reach the researchers (I've asked), but you could demand answers.

Law schools are essential institutions in our nation of laws. The barriers to admission are simply too arbitrary and inequitable. You have the power, responsibility and data to begin correcting that.

Please act on this now, so that a system you helped to put into place can be overhauled to the benefit of future generations of lawyers and our society as a whole.

10. EXHIBIT A: ERRORS AND PECULIAR CHOICES IN LSAT STUDIES

This section was moved from the main body of the communiqué for readability.

10.1. Introduction

Despite a solid background in statistics and stochastic processes, I do not offer what I recall (and have since refreshed) as a basis for credibility in those subject areas. My primary purpose, below, in pointing out what I believe to be statistical and measurement errors and peculiar choices is to stimulate discussion among qualified academics. The more they look at these studies, the more questions and concerns they will have about the statistical measures employed by LSAC, its methodology and, particularly, its analysis.¹¹³ Below, I merely scratch the surface of LSAC's pseudo-science.

10.2. A 30-second (visual) primer

Before we begin a discussion of the errors and peculiar choices, it will be helpful to show variants of some of the equations used by LSAC to perform its analysis of the LSAT. **No study or memorization is necessary by the reader!** A momentary appreciation of their complexity, particularly the fact that they are non-linear, is helpful when evaluating some of what LSAC attempts to demonstrate in its studies. For readability sake, we reduce the equations and brief discussion to footnotes for variance¹¹⁴, correlation coefficient¹¹⁵ (and, earlier, regression).

¹¹³ Some of this analysis will likely be proven incorrect by qualified statisticians. But, I trust that anyone disproving one argument will **not** be permitted to argue that everything here is somehow invalid or should be ignored. LSAC seems to have an unusual power to survive despite the criticisms lodged against it and the LSAT. Please do not allow them to discount the valid criticisms and analysis because they may discover and magnify imperfections in this communiqué.

¹¹⁴ See Wikipedia contributors, "Computational formula for the variance," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Computational_formula_for_the_variance (accessed approximately March 1, 2011). The variance of random variable x , or $\text{Var}(x)$, is a **theoretical** measure of how closely packed or scattered the possible values for x are. x could be LSAT scores or grade point averages, etc. Since we cannot measure variance directly, we define (an unbiased) sample variance, s^2 , or $\hat{\sigma}^2$ as:

$$\hat{\sigma}^2 \equiv \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{N}{N-1} \left(\frac{1}{N} \left(\sum_{i=1}^N x_i^2 \right) - \bar{x}^2 \right)$$

where x_i is an observed value of x and \bar{x} is the mean of the x values. $1/(1-N)$ is a factor that attempts to correct for the scenario in which the mean is determined from the same samples used to estimate the variance. When this is not the case or when the sample size is large, multiply the above by $(N-1)/N$. Note that when you perform scalar operations on a random variable (like LSAT score or FYA), the following occurs to the variance:

Multiplication by constant "c":	$\text{Var}(cX)$	$= c^2 * \text{Var}(X)$
Adding 2 random variables:	$\text{Var}(X + Y)$	$= \text{Var}(X) + \text{Var}(Y) - (\text{overlap in Variance})$

¹¹⁵ The quantity r , called the linear correlation coefficient (or Pearson product moment correlation coefficient), measures the strength and the direction of a linear relationship between two variables. r will always be between -1 and 1, it is defined as:

10.3. Analysis of Differential Prediction Studies – Race/Ethnicity

In the LSAC's latest national race and ethnicity LSAT vs. FYA study, *Analysis of Differential Prediction of Law School Performance by Racial/Ethnic Subgroups Based on 2005–2007 Entering Law School Classes (TR-09-02)*,¹¹⁶ the following errors and peculiar choices appeared (emphasis added):

- a. In attempting to compare the academic performance of racial/ethnic subgroups, as well as the predictive capabilities of the LSAT, the LSAC researchers chose to compare the simple arithmetic mean of various statistical measures for each subgroup (correlation coefficient, actual FYA, actual FYA – predicted FYA, etc.). However, none of the standard statistical tests for such comparisons, such as the *t*-test, were used.¹¹⁷ No reason for this was ever given. If the LSAC data violated assumptions of the *t*-test, then other alternatives, such as nonparametric methods¹¹⁸, were available.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

or

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{s_X} \right) \left(\frac{Y_i - \bar{Y}}{s_Y} \right)$$

where *n* is the number of data pairs in the sample being examined, (*X_i*, *Y_i*) are data pairs (such as LSAT score, actual FYA), \bar{X} is the mean of the *X* values, \bar{Y} is the mean of the *Y* values, *s_X* and *s_Y* are sample variances, calculated as shown above. “The Pearson correlation coefficient is invariant to separate changes in location and scale in the two variables. That is, we may transform *X* to *a* + *bX* and transform *Y* to *c* + *dY*, where *a*, *b*, *c*, and *d* are constants, without changing the correlation coefficient. Note, however, that more general linear transformations do change the correlation coefficient.” See Wikipedia contributors, “Pearson product-moment correlation coefficient,” *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient (accessed approximately March 1, 2011). However, there does not seem to be any interpretation in the literature for adding or averaging multiple correlation coefficients, each taken from a different sample or population.

¹¹⁶ See Norton, *et. al.*, *Analysis of Differential Prediction 09-02*.

¹¹⁷ See StatSoft, “Basic Statistics,” *StatSoft Electronic Statistics Textbook* (Tulsa, OK: StatSoft, Inc., 2011), <http://www.statsoft.com/textbook/basic-statistics/#Correlations> (accessed approximately March 1, 2011).

The *t*-test is the most commonly used method to evaluate the differences in means between two groups. For example, the *t*-test can be used to test for a difference in test scores between a group of patients who were given a drug and a control group who received a placebo. Theoretically, the *t*-test can be used even if the sample sizes are very small (e.g., as small as 10; some researchers claim that even smaller *n*'s are possible), as long as the variables are normally distributed within each group and the variation of scores in the two groups is not reliably different (see also Elementary Concepts). As mentioned before, the normality assumption can be evaluated by looking at the distribution of the data (via histograms) or by performing a normality test. The equality of variances assumption can be verified with the *F* test, or you can use the more robust Levene's test. If these conditions are not met, then you can evaluate the differences in means between two groups using one of the nonparametric alternatives to the *t*-test.

¹¹⁸ See StatSoft, “Nonparametric Statistics,” *StatSoft Electronic Statistics Textbook* (Tulsa, OK: StatSoft, Inc., 2011), <http://www.statsoft.com/textbook/nonparametric-statistics> (accessed approximately March 1, 2011).

However, the large sample size of the LSAT data would have made that decision extremely unlikely.¹¹⁹

That LSAC departed from standard statistical practice is not just suspicious, it effectively invalidates the study.

- b. According to the study, “Correlation values were *calculated separately* for each law school included in the study and *averaged across schools*.” (p.1) Correlation coefficients result from nonlinear calculations which are impervious to changes in the magnitude and uniform displacement of the underlying random variables. Simple linear averaging distorts the relationships between the underlying variables of each correlation coefficient, so that the unknown statistical measure that results is no longer a correlation coefficient and has dubious, if any, statistical value.
- c. According to the study, “The relationships between the predictor variables, LSAT and UGPA, and the criterion variable, FYA, are measured through the computation of correlation coefficients. ... For each school with a sufficient number of racial/ethnic minority students, *the correlation coefficients for each subgroup were calculated separately by law school and averaged for all schools included* in the comparison sample. Table 6 provides the correlations between the predictors—both alone and in combination—and FYA.” (p. 14) Since the study states that the actual FYA’s had been normalized for all law schools in the study, all of the students took LSAT exams and received scaled scores from what LSAC alleges is effectively the same distribution¹²⁰, and the undergraduate GPA’s were similarly scaled but in no way normalized within or across law schools, why would the researchers calculate correlations among the relatively tiny quantities of subpopulations at *each* school when they could have calculated far more accurate correlations for the four subpopulations by combining all data at all schools? Could it be that the actual FYA’s weren’t really normalized, in which case any comparison among schools is totally invalid? Did LSAC choose the approach most favorable to the LSAT?

Nonparametric methods are statistical procedures that enable us to process data of "low quality," from small samples, on variables about which nothing is known (concerning their distribution). Specifically, nonparametric methods were developed to be used in cases when the researcher knows nothing about the parameters of the variable of interest in the population ... nonparametric methods do not rely on the estimation of parameters (such as the mean or the standard deviation) describing the distribution of the variable of interest in the population.

¹¹⁹ StatSoft, “Nonparametric Statistics.”

Nonparametric methods are most appropriate when the sample sizes are small. When the data set is large (e.g., $n > 100$) it often makes little sense to use nonparametric statistics at all. ... In a nutshell, when the samples become very large, then the sample means will follow the normal distribution even if the respective variable is not normally distributed in the population, or is not measured very well. Thus, parametric methods, which are usually much more sensitive (i.e., have more statistical power) are in most cases appropriate for large samples.

¹²⁰ LSAC reports that the LSAT “curve” is determined from experimental sections of multiple prior tests, multiple test-takers.

- d. The study continues, “To further summarize these data, **the weighted average of the mean residuals between predicted and actual FYA** for each prediction equation/subgroup combination is provided in Table 7.” Regression equations rely on several sum of squares (SS) calculation, including the sum of the squares of residuals.¹²¹ The “mean residuals between predicted and actual FYA” do not appear in a regression equation, just the square of the residuals. This LSAC linear calculation is improper because any of these residuals can be negative, thus canceling other residuals out, artificially and falsely reducing evidence that these errors ever existed. The actual magnitudes of the means associated with residuals, therefore, could be much higher than reported. Why would LSAC chose a transformation that improperly diminished the errors in their analysis?
- e. According to the study, “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 alone, UGPA alone, and the combination of LSAT and UGPA. *Differences between predicted and actual FYA were then calculated for each subgroup* based on each regression equation. For the difference calculation, *the mean actual FYA earned by students at a participating school was subtracted from the mean predicted FYA for students* at the school. A *resulting negative difference indicates* that the regression equation underpredicted the mean (or average) performance of a subgroup in a law school, while a positive value indicates that the regression equation overpredicted the mean performance of a subgroup in a law school.” (p. 14)

There are levels upon levels of manipulations of approximations of non-linear

¹²¹ Investopedia Contributors, “Regression,” *Investopedia*, <http://www.investopedia.com/terms/r/regression.asp> (accessed approximately March 1, 2011). Regression takes a group of random variables, thought to be predicting Y, and tries to find a mathematical relationship between them. This relationship is typically in the form of a straight line (linear regression) that best approximates all the individual data points.

Linear Regression: $Y^{\wedge} = a + bX + u$

Multiple Regression: $Y^{\wedge} = a + b1X1 + b2X2 + B3X3 + \dots + BtXt + u$

Where:

Yi^{\wedge} = the variables that we are trying to predict

Xi = the variables that we are using to predict Y

(Yi, Xi) = the actual data pairs

u = the regression residual

From Wikipedia contributors, "Ordinary least squares," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Ordinary_least_squares (accessed approximately March 1, 2011), Ordinary least squares (OLS) or linear least squares is a method for estimating the unknown parameters in a linear regression model. It minimizes the sum of squared residuals (vertical distances between the observed responses in the dataset, Y_i , and the responses predicted by the linear approximation \hat{Y}_i) or $\sum (Y_i - \hat{Y}_i)^2$.

Summing **squared** residuals is critical so that the residuals never cancel each other out (Squares are always positive.) See also David W. Stockburger, “Regression Models,” *Introductory Statistics: Concepts, Models, and Applications*, <http://www.psychstat.missouristate.edu/introbook/sbk16m.htm#08> (accessed approximately March 1, 2011).

statistics as though they were scalars to be added, subtracted, averaged; subpopulation statistics were calculated using regression equations not solved specifically for that particular subgroup; the mean of predicted FYA's were calculated for subgroups, then further manipulated; conclusions were drawn about over- vs. under-prediction. It is doubtful that any meaningful analysis can come from these corruptions of linear approximations/snapshots of nonlinear statistics. And, yet, LSAC uses these custom manipulations to prove the underperformance of minority students?

- f. There would seem to be no methodologically sound reason that correlation coefficients were calculated *for each subgroup* independently at a given law school in one part of the study (**Predictive Validity**, p. 14) and, in another part of the study, calculated *for all students* at a university and only then applied to each subgroup in turn (**Predicting First-Year Averages**, p. 14). No explanation is offered. The results of the study would have to be internally inconsistent.

Similar errors occur in other studies of this series.

10.4. Predictive Validity of the LSAT - National

In LSAC's latest national study of the predictability of the LSAT, entitled *Predictive Validity of the LSAT: A National Summary of the 2007 and 2008 Correlation Studies*,¹²² the following errors and peculiar choices appear:

- a. The researchers followed the pattern established by numerous LSAC researchers and formed a summary statistic called the "mean" of the correlation coefficients which were measured for each university that participated in that study. Standard deviation calculations were also made of this soup of correlation coefficients. (See Tables 2 and 3, pp. 8 and 12, respectively.) However, the university means were taken from populations with different variances. When data points are derived from populations with different variances, they are termed heteroscedastic and special care must be taken when comparing them.¹²³ The researchers, however, apparently formed a simple arithmetic mean, introducing variance perturbations, at minimum. What is demonstrated when averaging a mean from a law school with 50 first-year students with the mean from one that contains 200 students, where the variances differ by 25%? The criticisms of averaging correlation coefficients stated above also apply. It is unclear, therefore, what this or any of LSAC's correlation studies actually measure. LSAC should have used standard statistical tests to compare the correlations.
- b. The researchers chose the term "selected group" to mean only those LSAT test-takers who actually matriculated and have first-year GPA's at a given university. The term "unselected group" means the entire set of applicants to a specific university *and this includes those who both matriculated and have first-year GPA's*. These two terms

¹²² See Stilwell, *Predictive Validity of the LSAT 09-03*.

¹²³ For a discussion of heteroscedasticity in regression, see Wikipedia contributors, "Heteroscedasticity," *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Heteroscedasticity> (accessed approximately March 1, 2011) or SAS Institute Inc., "SAS/ETS Examples: A Simple Regression Model with Correction of Heteroscedasticity," *SAS Knowledgebase Base*, <http://support.sas.com/rnd/app/examples/ets/hetero/index.htm> (accessed approximately March 1, 2011). The LSAC researchers form a simple summary statistic and apparently do not regress, but complications similar to regression errors could be introduced.

clearly overlap, while their names convey mutual exclusivity. This conflation creeps into the calculations, as in Table 3, which falsely claims that it presents “actual” correlation coefficients for the “unselected” population of applicants – an impossibility for the great majority of that class (who never matriculated).¹²⁴ That conflation also creates distortions in the attempts to compensate for the “selection effect” or “restriction of range problem.” The correlation coefficients in Table 3 should refer to an unrestricted population, not unselected. Though their definition of unselected seems to equate to unrestricted, the incorrect English usage of the former simply has to be red-flagged as indicative of deception there and possibly elsewhere, perhaps hidden in the calculations or unpublished data.

- c. In attempting to compensate for the “selection effect” or “restriction of range problem,” the researchers attempt to apply a technique sometimes known as “Thorndike’s Case 2” (in <http://pareonline.net/pdf/v14n5.pdf>, page 4). However, their formula in Appendix A would appear to be wrong for the bi-variate (2 variable) case, because it can never be negative. There are also questions about how accurate that transformation is.¹²⁵

Similar errors occur in other studies of this series.

10.5. LSAT Item-Type Validity

In the LSAC section/item type study quoted above *LSAT Item-Type Validity Study (TR 98-01)*,¹²⁶ the following errors and peculiar choices appear:

- a. The researchers used raw LSAT score instead of scaled score in their analysis because they falsely assert “[raw score] was the test score variable that was essentially being used in the selection process.” Universities always report scaled scores in public comparisons, not raw scores. (p. 3) Were the results more damaging to the LSAT if scaled scores had been used?
- b. The researchers stated that “For any applicant who took the LSAT more than once, use their most recent LSAT data.” (p. 3) However, numerous LSAC studies have indicated that the most accurate correlations are found when the average test score is used.¹²⁷ One must ask why they chose measures deemed inferior by LSAC.

¹²⁴ The specific text at Stilwell, *Predictive Validity of the LSAT 09-03*, 12 states “Table 3 presents estimated and **actual** correlation coefficients for the **unselected** population of applicants.” (emphasis added)

¹²⁵ See Wiberg and Sundström, “A Comparison.” Note that “Thorndike’s Case 2” is the less accurate of the two measures tested. It overstated the correction when direction is taken into account and understated it when only magnitude is considered. The standard deviations were 3 – 6 times greater in this Swedish study than in the LSAC study, however, so directly comparing the magnitude of the corrections is not helpful.

¹²⁶ See Louis A. Roussos, Lynne L. Norton, *LSAT Item-Type Validity Study, LSAT Technical Report 98-01* (Newton, PA: Law School Admission Council, March 1998), <http://www.lsac.org/LsacResources/Research/TR/TR-98-01.pdf> (accessed approximately March 1, 2011).

¹²⁷ Available before this study was completed was Susan P. Dalessandro and Lori D. McLeod, *Validity of Law School Admission Test Scores for Repeaters: A Replication, LSAT Technical Report 98-05* (Newton,

- c. There appear to be uncorrected heteroscedasticity and false assumptions of linearity, as discussed above.

10.6. General

Finally, it appears that all of the studies (available online) in the *repeat test-taker* series contain similar errors in calculating the aggregate subpopulation mean from among university means with different variances and fail to employ standard statistical tests to probe the validity of these comparisons. Similar errors, apparent oversights, and peculiar choices appear in most of the other studies I examined. Skillful researchers should be able to analyze them in greater depth and with greater precision.

PA: Law School Admission Council, September 1998),
<http://www.lsac.org/LsacResources/Research/TR/TR-98-05.pdf> (accessed approximately March 1, 2011).
It reports that averaging scores produces the most reliable results for repeat test-takers and that a previous study completed in 1990 found similarly. I choose not to indicate whether that appears backed by properly derived evidence, but merely report that, in the LSAC world, this was an accepted fact that was ignored by the researchers in question.