



THE UNIVERSITY OF ARIZONA
James E. Rogers
College of Law

DEAN'S OFFICE

James E Rogers College of Law
1201 E Speedway Blvd
PO Box 210176
Tucson AZ 85721-0176
520-621-1498 / law.arizona.edu

January 18, 2023

Mr. William Adams, Jr.
Managing Director
ABA Council on Legal Education &
Admission to the Bar
321 N. Clark Street
Chicago, IL 60654-7598

Dear Bill,

Thank you and the Council for commissioning an external consultant's review of the science supporting JD-Next as a valid and reliable test for predicting success in law school. We appreciate Dr. Kuncel's expert review of our body of research, consisting of multiple peer-reviewed studies reflecting years of work. We offer a few reflections to assist the Council in understanding our work and the cautions and caveats that Dr. Kuncel mentions.

ABA Standard 503 governs the Council's decision to approve the general use of the JD-Next final exam in law school admissions. Standard 503 provides:

"A law school shall require each applicant for admission as a first-year J.D. degree 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 program of legal education."

The question is whether JD-Next is a valid and reliable admissions test. Indeed, under Interpretation 503-1, a school using a test need only "demonstrate" it "is a valid and reliable test." Dr. Kuncel's clear bottom-line answer to the accreditation question is that:

"[b]ased on the evidence, I conclude that the JD-Next exam is a reliable and valid predictor of early law school grades..." (p. 8).

Dr. Kuncel based his conclusion on (1) the robust background literature supporting this type of test, (2) the fact that the exam covers material "directly relevant to the content of law school education," and (3) "the reliability and criterion related validity evidence from the 2019, 2020, and 2021 cohorts support the argument that the JD-Next exam has an acceptable level of reliability for making decisions and a useful level of predictive power for early law school grades" (p. 8, citations omitted).



As empirical researchers and legal scholars, we are devoted to evidence-based assessment. Therefore, in addressing Dr. Kuncel's additional remarks in his conclusions (p. 8), if the JD-Next final exam is approved for general use in law school admissions under ABA Standard 503, we remain committed to collecting additional JD-Next test and student outcomes data in future operational high-stakes admissions contexts. By collecting these data, we may continue to demonstrate the JD-Next final exam's strong validity and reliability.

We have five points to put the consultant's caveats in context.

First, validity is the ultimate psychometric concern. We agree with Dr. Kuncel that a longer test (i.e., a test with more questions) has higher psychometric reliability. Higher reliability is a means to achieve higher validity, but higher reliability does not necessarily result in higher validity. In fact, there can be diminishing returns for validity when one increases test reliability by increasing test length. The purpose of JD-Next and its final exam is to simulate a real law school environment, and the JD-Next final exam is the length of an authentic law school exam. Lengthening the JD-Next final exam would at a certain point result in the exam being unrepresentative of a real-world law school exam. Consistent with Dr. Kuncel's conclusions, *the existing levels of validity and reliability of the JD-Next final exam satisfy ABA Standard 503*. Indeed, the science has shown JD-Next validity to be comparable to legacy exams, such as the LSAT.

Second, aside from the evidence-based validity and reliability of the JD-Next test, our multiple peer-reviewed studies also found that *the test creates much smaller score disparities for historically underrepresented racial/ethnic populations, compared to the legacy law school exams*. Further, for the JD-Next exam, the differences in average scores between racial/ethnic groups were not statistically significant. As Dr. Kuncel correctly notes, this lack of significance should not be interpreted as a lack of a difference" (p. 6). Indeed, it is impossible to prove a negative -- that there is no difference in scores at all.

While Dr. Kuncel expresses the universal academic desire to have more data, our research uses the longstanding and settled approach for determining whether there is statistical significance, based on p-values at the .05 threshold and 95% confidence intervals.¹ Using these typical approaches, this very same data does confirm the large, statistically and practically significant disparities in legacy law school exams, consistent with standardized testing organizations' own published reports. This finding shows that we had statistical power to detect such effects and reject the hypothesis that the legacy exams avoid score disparities (see Figure 2, p. 157 in

¹ See Cumming G. Inference by eye: reading the overlap of independent confidence intervals. *Statistics in medicine*. 2009 Jan 30;28(2):205-20 ("when 95 percent confidence intervals (CIs) on independent means do not overlap, the two-tailed p-value is less than 0.05 and there is a statistically significant difference between the means.").

Findley et al., 2023). The confidence intervals also allow us to reject the hypothesis that JD-Next has disparities as large as the legacy standardized test.

Further, we find such legacy exam score disparities across three cohorts in two separate studies, and in all cases, we find that the JD-Next exam has significantly smaller score disparities by race. This replication helps to minimize the possibility that the first study's finding was due to random chance in a sample that was too small. Based on all these data, *the Council can be confident that the JD-Next exam is not only a valid admissions tool, but that it can reduce the exclusionary effects of law school admissions.*

Third, regarding the concern that the JD-Next exam may perform differently at schools with various levels of selectivity, our research investigated those possibilities by both using each student's school's median LSAT score as a proxy for its selectivity in multivariate regressions, and also testing for validity in subsets of schools grouped by their rankings. We were interested in "incremental validity," which measures whether the test provides more information, beyond the student's undergraduate GPA. *Across high-ranked, middle-ranked, and low-ranked schools across several cohorts, we consistently found that the JD-Next provided significant incremental validity.* Future research with additional data will employ multilevel modeling to further examine the issue of schools with different levels of selectivity.

Fourth, we offer two observations about Dr. Kuncel's concerns about range restriction, *i.e.*, that JD-Next has been evaluated in the research context of dozens of law schools, but may not cover the full range of potential test-takers in a high-stakes environment. There are statistical adjustments that can be made for range restriction, and they generally increase the observed correlations.² That is, if we had implemented these adjustments, or secured even broader samples, the observed correlations between JD-Next scores and first year grades would likely be even stronger, suggesting that it is even more valid as a predictor. Additionally, we note that *the best way to collect a broader range of test scores is to allow all law schools and all applicants to begin using the JD-Next exam in a high-stakes environment.*

Fifth, Dr. Kuncel's questions about test administration have largely been addressed by Aspen Publishing, the organization that has acquired an exclusive license to deliver the JD-Next exam. Dr. Kuncel expresses a question about test preparation and coaching, if "high quality

² See Schmidt, F. L., Oh, I. S., & Le, H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validities and other research results. *Personnel Psychology*, 59(2), 281-305 ("the size of effects and the strength of relationships are likely to have been substantially underestimated in any research literature where range restriction corrections are necessary"); Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychological bulletin*, 127(1), 162. ("Restriction of range results in underestimates of GRE validity coefficients for the actual applicant populations.") See *e.g.*, Wayne Camera, Evidence to Support Validity Claims for Using LSAT Scores in Law School Admission, *Law School Admissions Council* <https://www.lsac.org/data-research/research/evidence-support-validity-claims-using-lsat-scores-law-school-admission> (showing a raw correlation of .40 between LSAT and first year grades growing to .60 when adjusted for range restriction).

preparation [were] only available to those with the means to pay, the predictive power and fairness of the assessment may be compromised” (p7).

Unlike legacy exams, *the JD-Next exam is designed around a course, which has also been shown to boost law school performance*. Aspen has chosen to offer the JD-Next course bundled with the JD-Next exam for a single reasonable price, so all test-takers can benefit from the state-of-the-art pedagogical experience. In this way, we level the playing field for all test-takers, as Dr. Kuncel suggests.

Dr. Kuncel also advises that, when offered at scale, JD-Next needs to implement multiple forms, test security, and disability accommodations. Aspen has embraced all these responsibilities, to ensure a credible test.

~ ~ ~

Ultimately, Dr. Kuncel agrees that JD-Next is a valid and reliable test for law school admissions. That is what Standard 503 requires.

We look forward to continuing the research agenda to progressively address the caveats and cautions, following the science-based approach that created JD-Next over the past five years. We are also pleased that Aspen Publishing is administering the course and exam in a responsible way to ensure continuing validity and equity in a high-stakes environment.

Please let us know if the Council has additional questions.

Sincerely,

Christopher Robertson, JD, PhD
Founding Principal Investigator

Marc Miller, JD
Co-Principal Investigator
Dean, University of Arizona James E. Rogers College of Law

Jessica D. Findley, JD, PhD
Principal Investigator

MLM/mws

