A Case Report of Holistic Review of Graduate Applications

Heather Stoll, *University of California, Santa Barbara, USA*
Bruce Bimber, *University of California, Santa Barbara, USA*

In Fall 2020, with the pandemic and racial justice protests that were part of the Black Lives Matter movement vivid in our minds, our department's graduate admissions committee overhauled its process. We set aside previous admissions practices and applied a holistic-review process that was new to us, with the goal of achieving greater diversity in our graduate student body. This article reports on what we did and why. It provides an overview of holistic review, contrasting it with traditional graduate admissions approaches, and identifies the key questions and issues that those potentially interested in experimenting with holistic review must consider. Although there is not yet sufficient data to make evidence-based choices about holistic review, we are cautiously optimistic based on our experiences and the literature. Our goal in discussing our experiences in this article after only one admissions cycle—too soon to test for effects quantitatively in a program with small numbers—is to encourage conversations, experimentation, and future research about diversity and graduate admissions in political science.

THE PROBLEM

We think about diversity from the perspective of the "leaky pipeline," in which a critical leak is known to occur between undergraduate and graduate education. In the political science profession as a whole, minorities earned 40% of undergraduate degrees but comprised only 29% of incoming PhD students. By comparison, women earned 48% of undergraduate degrees and comprised 45% of incoming PhD students.[1] The situation in our department at the University of California, Santa Barbara (UCSB) aligns with this pattern. About 41% of our undergraduate majors are underrepresented minorities, defined by the University of California (UC) as those who identify as Black, Latino/Hispanic, and Native American/Alaskan Native.[2] Yet, in recent years, only about 26% of those admitted to our PhD program have been underrepresented minorities, with great variance among years. We lack data at the graduate level on first-generation status, but we expect that the proportion is much lower than the 44% of our undergraduate majors who are first generation. These numbers frame our challenge for improving diversity. Doing so is especially important because UCSB is both a Hispanic-Serving Institution and an Asian American Native American Pacific Islander Serving Institution.

INTRODUCTION TO OUR REFORMS

We undertook the overhaul of our admissions process with the short-term goal of reducing the undergraduate-to-graduate pipeline leak for the class entering in Fall 2021 and the long-term goal of a graduate student population that resembles our undergraduate student body. Our approach was informed by the principle of holistic review, which is "a process by which programs consider a broad range of characteristics, including noncognitive and personal attributes, when reviewing applications for admissions" (Kent and McCarthy 2016, 1). Holistic review is "widely viewed as a useful strategy for improving diversity of higher education" and is practiced in many settings, including undergraduate and medical school admissions (Kent and McCarthy 2016, 1). Although it has its critics, we were persuaded that holistic review has received enough support for us to adopt it in the spirit of experimentation. We intend to build a multiyear body of evidence about its effects.

Our reforms were guided by Kent and McCarthy's (2016) review of holistic graduate admissions as well as by the general review of evaluation and decision making in higher education by Posselt et al. (2020). We additionally drew from descriptions of holistic review developed outside of political science, especially the highly influential Fisk–Vanderbilt Bridge Program (Stassun et al. 2011).[3] Some departments at UCSB use holistic review, and our institution provides National Science Foundation (NSF)–supported guidance about making the change.[4] This meant that we had institutional support, examples of how others achieved it, and good reasons to believe that our traditional process should be updated.

TRADITIONAL GRADUATE ADMISSIONS IN POLITICAL SCIENCE

Our previous practice was to rate applicants on a single evaluative dimension that frankly was not well defined. Although admissions committee members in recent years universally voiced support for diversity, equity, and inclusion, our process did not formally incorporate these considerations. Admissions committee members assigned a score to each applicant using their individual preferences for how to weigh the various indicators in the written file—for example, test scores, Grade Point Average (GPA) and relevant coursework, references, status of undergraduate or graduate institutions, personal statement, writing sample, and any reported contributions to diversity. In addition to being dissatisfied with

doi:10.1017/S1049096522000865
© The Author(s), 2022. Published by Cambridge University Press on behalf of the American Political Science Association.

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press

diversity outcomes from this approach, we were skeptical of the reliability and validity of the scoring.

Based on our informal review of the information displayed on their websites, our previous approach appears consistent with those of many other political science departments.[5] Among the top 25 departments, six (23%) did not disclose any information about how admissions decisions are made or which evaluative criteria are used. Of the remaining 20 departments (77%) that provided information, all described use of a comprehensive process—for example, no specific weighting of Graduate Record Exam (GRE) scores or minimum GPA, careful review of all components of the application, and no formula—that resembles holistic review. However, only nine departments (35%) stated that they consider diversity, personal characteristics, or other noncognitive attributes. Moreover, only one department described a process that, in our view, qualifies as holistic through the use of specific, stated evaluative criteria that surpass traditional cognitive indicators.

Although it has its critics, we were persuaded that holistic review has received enough support for us to adopt it in the spirit of experimentation.

HOLISTIC REVIEW: KEY PARAMETERS

Many different practices are associated with the term "holistic review." This section outlines the key parameters that we identified from the resources discussed previously, the decisions that we made and our reasoning for them, and alternatives.

Use of the Graduate Record Exam

The first decision we made was whether to require the GRE. As in holistic review more generally, this decision was subject to debate. Our review showed that 10 (38%) of top departments currently do not require GRE scores. Neither the 10-campus UC system as a whole nor our own campus has suspended use of the GRE as an institutional decision; however, the UC system has suspended the SAT and ACT for undergraduate admissions. Other departments on our campus decided against the GRE, and we did so as well. One factor in our decision making was the pandemic, given the difficulty that some prospective applicants experienced in taking the exam. However, also driving our decision was the growing body of research suggesting that the GRE is biased against women and minorities, as well as that among those who enroll, it has limited correlation with student success, defined in various ways such as completion of the PhD (Miller and Stassun 2014; Miller et al. 2019; Stenberg and Williams 1997; Wilson et al. 2019). However, the case against the GRE is far from settled.[6] We are not aware of any empirical work on its predictive role in political science admissions traditionally[7] or on exactly how much information the GRE adds to holistic-review processes.[8] Contrary to our approach, it is plausible to use the GRE in a holistic-review process as long as it is thoughtfully contextualized to avoid known pitfalls, such as explicit or implicit cutoffs (Kent and McCarthy 2016, vi, 12, 15–17; Wilson et al. 2019).

First-Round Screening Criteria

The next decision we made was whether to use a first round of screening before applying the full set of holistic review criteria. We expected (correctly) that holistic review would require considerably more time than our traditional review, especially with interviews included. Indeed, the time-intensive nature of holistic review is generally viewed as one of the major barriers to its use (Kent and McCarthy 2016). We decided to reduce our full set of applicants (N=127) to a subset (N=30) of quarter-finalists more workable for our four-person admissions committee. Annual variation in funding and target cohort size, committee size, and the period for completing reviews will shape future screening targets.

We settled on a primary screening criterion of fit between the applicant's areas of interest and the expertise of our faculty. As a medium-sized department, we do not have a deep roster in every area of the discipline; therefore, we screened out applicants whose interests did not align with the specific expertise and active research programs of at least two faculty members. We consider fit to be a good initial screening criterion for larger departments as well. Establishing criteria for the initial screening was a challenge because we needed to avoid slipping into our process the traditional cognitive indicators that we were trying to keep in perspective: quantitative measures such as GPA and shortcut qualitative metrics such as the prestige of the undergraduate institution. These indicators often are used for initial screening (Kent and McCarthy 2016, 15–17) but are known to privilege groups that are overrepresented (Posselt et al. 2020).

We also used a second criterion that proved unwieldy: basic readiness for doctoral study. This was difficult to implement and did not screen out many applicants. In the future, we either will omit or significantly rework this criterion.

Our third element of screening is one that we may not repeat. We contacted prospective faculty advisors for the 30 quarter-finalists and asked them to provide a simple yes or no signal of their willingness to take on the applicant as an advisee if admitted. This step led to rejecting more applicants than we had anticipated; the result was a set of 20 semi-finalists who would proceed to the full holistic review.

We have concerns about the advisor screening step. Some faculty may have responded on the basis of the type of traditional criteria that we were attempting to keep in perspective with holistic review. In the future, we either will provide faculty with explicit guidance about the evaluation criteria or eliminate this step altogether.

Overall, we were struck by how challenging it is to screen a large number of applicants to attain a set that, given workload

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press

considerations, is manageable for holistic review without defeating the purpose. Given sufficient faculty time, it would be ideal to use a holistic-review process for all applicants without initial screening.

Second-Round Holistic Review

This subsection describes the key decisions we made with respect to the holistic review of the screened applicants.

Evaluation Criteria and Rubric

Our next two decisions concerned which evaluation criteria to use and the associated rubric. Even departments that are interested only in evaluating cognitive criteria along traditional lines should benefit from this step due to its potential for minimizing bias and making the evaluation process more transparent, consistent, and efficient (Kent and McCarthy 2016, vi, 13). To evaluate the semi-finalist applicants, we adopted five dimensions of evaluation, including elements related to diversity and "noncognitive" qualities relating to adjustment, motivation, and perceptions (Kent and McCarthy 2016). These five dimensions were academic preparation, relevant research experience, perseverance, self-knowledge, and leadership-outreach activities/DEI. We drew heavily on the criteria used by existing, well-regarded programs with an established diversity-promoting record, notably the Fisk–Vanderbilt Bridge Program (Stassun et al. 2011).[9] Other criteria are also in use around the country. For example, the UC–San Diego political science department uses "diversity" as the sole noncognitive criterion along with otherwise more conventional cognitive criteria.[10] Solely cognitive criteria also could be used. Which criteria are the most helpful in delivering "inclusive excellence," to use Posselt's (2013) phrase, is a question that unfortunately is well ahead of evidence at this point.[11] Yet, the same can be claimed about traditional approaches to admissions that examine multiple indicators of academic performance. We believe that the discipline needs experience gained from programmatic experimentation and data collection.

We scored each applicant as high, medium, or low on each of the five criteria using a rubric that we adapted from the sources cited previously (figure 1). We wanted to minimize implicit bias and values, especially those associating quality with prestige (Posselt 2013), and also to avoid overweighing what can be quantified (e.g., GPA). We experienced a tendency toward rating at the mean, but we were mostly satisfied with the process. In the future, we will revise the rubric to produce more spread and we will engage in an initial round of norming. We also welcome and anticipate further discussion of appropriate criteria and their operationalization.

Interviews

Next, we decided whether to conduct interviews. We chose to interview all semi-finalists using a written interview protocol derived from the Fisk–Vanderbilt Bridge Program. Some of our questions are listed in figure 2. These interviews provided useful information in addition to what was in the written files, especially for the noncognitive criteria. Because of how much light our conversations shed on applicants' qualities beyond their academic record, we will continue to interview in the future. After the interviews, one committee member expressed amazement that the department had ever managed to admit people without having spoken with them first.

For other departments considering the adoption of interviews, we recommend discussing the risk of introducing a new avenue of bias and how this might be mitigated, such as by using a standardized set of questions. We were aware of this risk, but we believed that the importance of learning more about the applicants' noncognitive qualities would outweigh the risk of implicit or explicit bias in the interview. An alternative for those who are concerned about interview bias involves application instructions. We encourage departments to direct applicants and their letter writers to address whatever evaluation criteria are being used, especially noncognitive criteria. This is especially useful for letter writers to understand that the admissions committee is requesting more than a traditional assessment of academic performance. Due to the timeline for developing our new process, we could not communicate our holistic-review criteria in advance to candidates and letter writers, but we will do so in future admissions cycles. If adequate information can be elicited in the written materials, interviews might have less importance.

Score Aggregation

After the applicants were rated, informed by our interviews, our final step was to explore approaches to aggregating each semi-finalist's five scores. These approaches ranged from simple and weighted averages to a count of the number of criteria on which the applicant scored "high." Of course, many alternatives can be imagined. In our case, these approaches converged fairly well, mostly all pointing to the same set of semi-finalists, who we then divided into finalists: admitted and waitlisted candidates. In a case in which various approaches to aggregating scores do not converge, admissions committees must apply another set of principles in this final stage.

Our next two decisions concerned which evaluation criteria to use and the associated rubric. Even departments that are interested only in evaluating cognitive criteria along traditional lines should benefit from this step due to its potential for minimizing bias and making the evaluation process more transparent, consistent, and efficient.

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press

Figure 1

Rubric for Second-Round Holistic Review

	High	Medium	Low
Academic Preparation	Really strong GPA, such as A- or A; no concerning grades (less than B) in key courses; at least one academic honor or advanced degree; undergrad or grad degree in political science or closely related discipline (e.g., international affairs, global studies, public administration) [or a discipline relevant to goals, such as enviornmental science or economics]; language skills relevant to interests. Experience at a US university or comparable top university where US-style social science is practiced	Good but not outstanding GPA, such as B+/A-; no concerning grades (less than B, especially in key courses) have reasonable explanation; some evidence of improvement; undergrad or grad degree in political science or closely related discipline (e.g., international affairs, global studies, public administration) [or a discipline relevant to goals, such as envirornmental science or economics]	Weak GPA (Note that the UCSB minium for grad applicants is B/3.0, so that is the lowest GPA we should see.) Concerning grades do not have a reasonable explanation; little evidence of improvement; degree in field unrelated to political science or area interests
Relevant Research Experience	Has engaged in and learned from significant research experiences outside of the classroom at a high level, such as an academic publication or professional conference presentation, or research experience equivalent to a senior thesis (senior thesis, intensive multi-month independent or work-related research project, master's thesis); has technical skills relevant to interests, such as statistics, interviewing, data collection of some kind; can succinctly articulate political science research interests (clear about details and big-picture implications)	Shows some evidence of research experiences outside of the classroom, and learning from them: at least one academic publication or conference presentation in progress, or some related publication or presentation at the campus/non-professional level; some research experience working as a research assistant or undertaking a research project less than a senior thesis, where some learning is reported; some difficulty succinctly articulating political science researchinterests (vague/general or not squarely political science)	Has not engaged in research outside of the classroom: no academic publications or conference presentations in progress; no research experience; unable to articulate political science research interests
Perseverance	Can describe a time they failed or encountered an obstacle and successfully coped, especially ongoing challenges or obstacles; communicated long-range goals beyond the PhD	Can identify a time they hit an obstacle but has trouble defining how they overcame the challenge; primary goal is PhD	Has little experience with failure/obstacles; cannot provide an example or describe a response; vague about long-range goals, or goals are short-term such as coursework
Leadership/ Outreach Activities	Demonstrates involvement and leadership ability in either academics, family, community, religious groups, or athletics; has been or has plan for being an advocate for DEI in some way	Demonstrates involvement in groups in academia or extramural, but has not shown leadership; some evidence of engagement with DEI	Not involved in groups in academia or extramural; no leadership; limited or no engagement with DEI
Self-Knowledge	Can clearly and realistically delineate strengths and weaknesses; works on self-development; expresses confidence they can complete goals; makes positive statements about abilities; can identify a professional support network including mentors	Has trouble identifying strengths and weakness but appreciates/seeks both positive and negative feedback; shows confidence and independence but may be unsure about adequacy or skills; expresses support from one individual, family, or community	Over or understates abilities; does little to no self-assessment; does not appear to have learned from experiences; is unsure they can complete the program/exhibits low self-esteem; expresses little or not support from family or community

Adapted from Stassun et al. (2011).

OUTCOMES

We were extremely pleased with the final list of admitted and waitlisted candidates, and we were sufficiently satisfied with the process to continue gaining experience with it. Our top candidates evidenced the type of noncognitive qualities that we anticipated were associated with successful completion of a doctoral program, as well as academic excellence. The diversity of the finalists exceeded that of our existing graduate population, which meant that we met our short-term goal. Specifically, a large majority (73%) of the finalists were women; a solid majority (64%) were underrepresented minorities; slightly more than a quarter (27%) were racial or ethnic minorities

Our top candidates evidenced the type of noncognitive qualities that we anticipated were associated with successful completion of a doctoral program, as well as academic excellence. The diversity of the finalists exceeded that of our existing graduate population, which meant that we met our short-term goal.

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press

Figure 2

Sample Interview Questions

Give an example of a time you hit a serious obstacle or failure, inside or outside of your academic career, and how you handled it. [Perseverance]

We are interested in whether you have had to deal with any systematic disadvantages or challenges in your life. If so, can you tell us about that and how you dealt with it? (for clarification if the interviewee asks: E.g. Some first-generation college students face particular obstacles, as do people who are subject to discrimination of various kinds.) [Leadership/Outreach/DEI]

Follow up: Have you done anything to serve as a role-model or leader for others facing similar challenges?

What motivates you to get a PhD in political science and what do you want to do with it? [Perseverance, Leadership/Outreach/DEI]

What do you do when you get in over your head? [Perseverance, Self-Knowledge]

What extracurricular or community activities are you engaged in? [Leadership/Outreach/DEI]

Follow up: Have you played a leadership role? How did things come about, and what did you do?

Follow up: Have you advocated for others in some way?

Adapted from Stassun et al. (2011).

not classified as underrepresented; and slightly more than a third (36%) were first-generation students. Also, a substantial number were in the intersectional category of underrepresented minority women.

CONCLUSION

This article briefly describes traditional admissions procedures in political science doctoral programs, identifies key issues involved in a holistic-review process, and recounts our department's experience with this process in the 2020–2021 cycle. Our experience, along with the literature on graduate admissions, leads us to recommend that other departments (further) explore holistic review, moving beyond known problems of bias and lack of transparency in traditional admissions as well as compiling evidence about outcomes that can inform future improvements. The primary motivation for our own choice to modify admissions was to reduce the leakage of minorities in the undergraduate–graduate pipeline in political science and, in this way, to further diversity in the discipline. Other departments may have similar or different motivations. We emphasize that our experience does not constitute definitive evidence in support of holistic review or particular ways of accomplishing it. We look forward to the political science discipline building a body of research evidence about graduate admissions processes, especially about the role of the GRE, measurement in general and in assessment of noncognitive attributes in particular, and how different evaluative criteria relate to different types of success. Our hope is that this case report will encourage such research and contribute to a much-needed conversation about diversity and admissions in political science doctoral programs.

ACKNOWLEDGMENTS

The authors are two of a four-member admissions committee that collectively made the decisions described here. We acknowledge the other two members, Satyajit Singh and Tae-Yeoun Keum, who participated equally in the study. We also thank Department Chair, Kate Bruhn; Associate Dean of Graduate Division, Paige Digeser; and Associate Dean–Faculty Equity Adviser of the Bren School, Sarah Anderson, for their support of and assistance with this initiative.

CONFLICTS OF INTEREST

The authors declare that there are no ethical issues or conflicts of interest in this research. ■

NOTES

1. The 2018–2019 data are from APSA's P-WAM dataset (American Political Science Association 2020).
2. All data on political science undergraduate majors are from the Fall 2018 Undergraduate Enrollment Profile for Political Science, produced by UCSB's Institutional Research, Planning & Assessment Office. Data on graduate students are from UCSB's Diversity Programs Office in Graduate Division and cover admissions from 2017 to 2019.
3. We replied on their publicly available toolkit, developed by Hall, Arnett, Cliffel, Burger, and Stassun (2011) (www.nsf.gov/awardsearch/showAward?AWD_ID=1110924) with funding from NSF Grant No. HRD-1110924 (www.fisk-vanderbilt-bridge.org/toolkit).
4. More information about this initiative is available at pullias.usc.edu/c-cide. It is funded by National Science Foundation Innovations in Graduate Education Grant No. DGE-1807047.
5. This survey was conducted in September 2021, when department websites were oriented toward the 2021–2022 admissions cycle. Rankings are from the 2021 *U.S. News & World Report*. Due to a tie at 25, 26 departments were surveyed.
6. See, for example, Kuncel and Hezlett (2010), who reported evidence that the GRE is predictive of outcomes such as degree attainment and citation counts, and they argued against bias.
7. The one partial exception is King, Bruce, and Gilligan (1993). However, in addition to being almost 30 years old, this study of the graduate-admissions process at Harvard did not parse out the independent effect of the GRE from other quantitative factors (e.g., GPA).
8. For some steps in this direction, see Wilson et al. (2019).
9. We renamed the Fisk–Vanderbilt "Communication/organizational ability/maturity/collaboration criterion as "self-knowledge" but did not modify the content.
10. These criteria, appearing as part of the "Admissions FAQs" on the department's website, are evidence of potential for research skills, writing and

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press

communication skills, analytical-thinking skills, contributions to diversity, and academic effort and preparation.

11. As a case in point, consider the lack of definitive recommendations in Kent and McCarthy (2016).

REFERENCES

American Political Science Association. 2020. "Project on Women and Minorities Dashboard (P-WAM)." Dataset generated February 2020. https://apsanet.org/resources/data-on-the-profession/dashboards/p-wam.

Kent, Julia D., and Maureen T. McCarthy. 2016. "Holistic Review in Graduate Admissions: A Report from the Council of Graduate Schools." Washington, DC: Council of Graduate Schools.

King, Gary, John M. Bruce, and Michael Gilligan. 1993. "The Science of Political Science Graduate Admissions." *PS: Political Science & Politics* 26 (4): 772–78.

Kuncel, Nathan R., and Sarah A. Hezlett. 2010. "Fact and Fiction in Cognitive Ability Testing for Admissions and Hiring Decisions." *Current Directions in Psychological Science* 19 (6): 339–45.

Miller, Casey, and Keivan Stassun. 2014. "A Test That Fails: A Standard Test for Admission to Graduate School Misses Potential Winners." *Nature* 510: 303–304.

Miller, Casey, Benjamin M. Zwickl, Julie R. Posselt, Rachel T. Silvestrini, and Theodore Hodapp. 2019. "Typical Physics Ph.D. Admissions Criteria Limit Access to Underrepresented Groups but Fail to Predict Doctoral Completion." *Science Advances* 5 (1). https://doi.org/10.1126/sciadv.aat7550.

Posselt, Julie R. 2013. "Towards Inclusive Excellence in Graduate Education: Constructing Merit and Diversity in Ph.D. Admissions." *Journal of Higher Education* 120 (4): 481–514.

Posselt, Julie R., Theresa E. Hernandez, Cynthia D. Villareal, Aireale J. Rodgers, and Lauren N. Irwin. 2020. "Evaluation and Decision Making in Higher Education: Toward Equitable Repertoires of Faculty Practice." *Higher Education: Handbook of Theory and Research* 31:1–63.

Stassun, Keivan G., Susan Sturm, Kelley Holley-Bockelmann, Arnold Burger, David J. Ernst, and Donna Webb. 2011. "The Fisk–Vanderbilt Master's-to-PhD Bridge Program: Recognizing, Enlisting, and Cultivating Unrealized or Unrecognized Potential in Underrepresented Minority Students." *American Journal of Physics* 79:374–79.

Sternberg, Robert J., and Wendy M. Williams. 1997. "Does the Graduate Record Examination Predict Meaningful Success in the Graduate Training of Psychologists? A Case Study." *American Psychologist* 52 (6): 630–41.

Wilson, Marenda A., Max A. Odem, Taylor Walters, Anthony L. DePass, and Andrew J. Bean. 2019. "A Model for Holistic Review in Graduate Admissions That Decouples the GRE from Race, Ethnicity, and Gender." *CBE—Life Sciences Education* 18 (1): ar7.

https://doi.org/10.1017/S1049096522000865 Published online by Cambridge University Press