Holistic Review for Graduate Admissions & Faculty Recruitment

Julie R. Posselt
University of Southern California
jrposselt@gmail.com
Prabhat: He grew up in a yurt in the Himalayas, was raised by his mom and grandma after his father died at an early age, and the next neighbors were two mountains over. He then found his way to a major U.S. public research university and has since started the only organization for the discipline in the Himalayan region.

Jeff: But do we think he can succeed?

[long pause]

Prabhat: He’s the most amazing case we’ve ever seen.

George: He would bring some personality to the department. I commit to look after him…. He presents himself as quite intelligent.

Chris: Excellent idea to give him a chance.

[Student ultimately admitted and enrolled.]
Why holistic review?

1) Talent is everywhere! Holistic review is a primary strategy for making excellence inclusive.

2) Individual criteria are...
   - ...ineffective indicators of the complex qualities we look for in students.
   - ...not equally distributed across sub-populations.

Therefore, we need to systematically collect multiple types of information about applicants, review them according to multiple criteria & in context.
Common evaluation & selection practices
Evaluation permeates academia.

<table>
<thead>
<tr>
<th>Production &amp; reception of academic work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition of scholars</td>
</tr>
<tr>
<td>Status of academic entities (universities, programs, journals, etc.)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admitted</td>
</tr>
<tr>
<td>Hired</td>
</tr>
<tr>
<td>Tenured</td>
</tr>
</tbody>
</table>
Ad hoc evaluation

• Responding to email inquiries from prospective students
• Reviewing letters of recommendation
• Reviewing writing samples
• Hiring postdoctoral fellows

Formal evaluation

• Peer review
• Book reviews
• Writing letters of recommendation
• Promotion/tenure review
• Annual awards, grant & fellowship funding panels
Review of student applications is a hybrid of ad hoc & formal evaluation.
"We are **not able to accept** students currently enrolled at the University of Michigan.

- **We seek** rising seniors, bachelor’s degree holders, and those currently in terminal M.A. programs.

- Please note, those applying to enroll in a doctoral program for Fall 2018 are **ineligible to participate**.

- **We seek** students who come from an educational, cultural, or geographic background that is underrepresented in graduate study in their discipline in the United States or at the University of Michigan.

- Students interested in **any field in the humanities** are eligible.

- To apply, **students should have a compelling record** of academic accomplishments.”
“Your application must include:

• A brief **essay** (approximately 500 words) on your academic and professional background and goals.

• **Unofficial transcripts** reflecting all undergraduate and graduate work (if applicable) through December 2017.

• Two **letters of recommendation**.

• An academic **writing sample** (between 7 to 15 pages double spaced). This paper should be a piece of original scholarship.”
Centrality of evaluation to academic life

• Academic disciplines are “status economies.”
• Recognition is among the most important currencies.
• Evaluation is a practice of patrolling our boundaries
  • Through it, we informally define who and what work is central/marginal.
    • Which topics are more/less likely to show up in “top” journals?
    • Who/whose work at is deserving of endowed professorships?
Analytic perspectives on academic evaluation

• **Functionalist:** How evaluation serves its purposes
  • Select individuals on the criteria that best predict success.

• **Power-analytic:** How evaluation dysfunctions create and maintain structural inequalities
  • Avoid criteria that are inequitably distributed by race, class, gender, etc.

• **Performativity:** How evaluations trigger organizational and individual behavior
  • Ranking algorithms discipline universities and programs to prioritize certain qualities in prospective members.

• **Constructivist:** How evaluation creates and reflects cultural values.
In principle, scholarly merit is the basis for access & advancement.

In practice, scholarly merit is something that we create when we make evaluations. It reflects and reinforces our values.

- Fellowship review derives from “disciplinary styles” (Lamont, 2009).
- Editorial judgments depend on the ”intellectual milieu” (Hirschauer, 2010).
- Academic book publishing is driven by “market logics” (Powell, 1985).
- Graduate admissions is shaped by “disciplinary logics” (Posselt, 2016)
  - High-consensus fields
  - Epistemology, theories, metaphors/images, and practical priorities of the field shape
  - Criteria in use
  - Beliefs about intelligence
  - Perceptions of a legitimate selection process
Why diversity in graduate education?

Moral & social good
- Representation similar to the population is a signal of equity
- To reduce inequality in the labor market

Educational benefits
- Non-divisiveness amid differences
- Cognitive complexity
- More likely to graduate

Business case
- Diversity helps UG recruitment & rankings
- Expands the technical workforce & middle class
- Practical benefits of diverse scientific teams

Civic development

To reduce inequality in the labor market
Common admissions practices in large, highly selective PhD programs
• **Research Questions:**
  • How do faculty individually judge & collectively select applicants to highly ranked Ph.D. programs?
  • What assumptions about merit guide faculty judgment
  • How do disciplinary norms shape faculty judgment?

• Comparative ethnographic case study
• 10 programs in 3 public & private universities
  • 85 interviews with professors & a few graduate students
  • 22 hours of admissions meeting observations in six of the programs
# Programs Studied

<table>
<thead>
<tr>
<th></th>
<th>Humanities</th>
<th>Social Sciences</th>
<th>Natural Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Consensus</strong></td>
<td>Philosophy (2 programs)</td>
<td>Economics</td>
<td>Physics</td>
</tr>
<tr>
<td><strong>Moderate Consensus</strong></td>
<td>Classics</td>
<td>Sociology</td>
<td>Astrophysics</td>
</tr>
<tr>
<td><strong>Low Consensus</strong></td>
<td>Linguistics</td>
<td>Political Science</td>
<td>Biology</td>
</tr>
</tbody>
</table>
Evaluative cultures explain apparent tensions between definitions of merit & valuing diversity.

- *Preference for specific criteria* was rooted in beliefs about what they signal. Those beliefs relate to their roles as scholars in highly ranked programs.

- *Preference for a process that is efficient and collegial.* Goals: Quantify quality & minimize conflict.

- In high-consensus fields like physics, *shared disciplinary norms* shaped working definitions of “merit”, ideas about intelligence & what the admissions process should look like.

- In low-consensus fields like political science and linguistics, *individual preferences* were as important as shared preferences in high-consensus fields and reflected patterns of homophily (“love of the same”).

- *Ambivalence about organizational change*, especially reforms related to diversity and equity.
Two-tiered review is used in most places.

<table>
<thead>
<tr>
<th>Initial screening</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conceptualizing merit</strong></td>
</tr>
<tr>
<td>Conventional achievers with low perceived risk of attrition</td>
</tr>
<tr>
<td><strong>Important criteria</strong></td>
</tr>
<tr>
<td>“Numbers” in context of undergraduate prestige &amp; curriculum rigor</td>
</tr>
<tr>
<td><strong>Relationship of merit &amp; diversity</strong></td>
</tr>
<tr>
<td>Standard of merit may be in tension with racial/gender diversity aims.</td>
</tr>
</tbody>
</table>
Two-tiered review is used in most places.

<table>
<thead>
<tr>
<th>Conceptualizing merit</th>
<th>Initial screening</th>
<th>Later rounds of review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptualizing merit</td>
<td>Conventional achievers with low perceived risk of attrition</td>
<td>Future of the discipline</td>
</tr>
<tr>
<td>Important criteria</td>
<td>“Numbers” in context of undergraduate prestige &amp; curriculum rigor</td>
<td>Experience with and dispositions for research; Unique perspective; research interests align</td>
</tr>
<tr>
<td>Relationship of merit &amp; diversity</td>
<td>Standard of merit may be in tension with racial/gender diversity aims.</td>
<td>Diversity is a component of merit.</td>
</tr>
</tbody>
</table>
Problems with the typical approach

Overreliance on metrics (without considering context) disproportionately excludes already underrepresented populations

Limited efficacy

Blind spots
From ETS document
"Factors that can influence performance on the GRE general test 2006-2007"
What does the literature say about GREs’ ability to predict student success?

Meta-analyses come to differing conclusions about the GRE’s validity.

- Morrison & Morrison, 1995;
- Kuncel, et al., 2001;
- Kuncel & Hezlett, 2010
- Orlando, 2005

Why?

Studies draw upon different methods, different disciplinary and institutional contexts, and different populations.

Only a few correct for attenuation bias; ETS continues to revise the test.

What do we know?

- There has never been a true validity study conducted: denied students aren’t studied.
- Correlations vary by exam and by graduate school outcome (Kuncel & Hezlett, 2007).
- The longer the time between the test and the outcome, the weaker the validity.
- A flurry of discipline-specific studies: some find relationships with first year graduate school GPA, none with later outcomes, race or gender (despite score gaps)
  - Psychology: Sternberg & Williams (1997)
  - Marine Sciences: Dore, 2017
  - Biomedical Sciences: Moneta-Koehler, et al., 2017; Hall et al., 2017
  - Physics: Miller et al., 2018
Limited Significance in Predicting Success

Miller et al., submitted
Admissions is only one reason for low completion rates. Peer mentoring and progress monitoring are also critical factors in retention.
The usual weight given to GRE scores exceeds its predictive capabilities and has negative societal impact.
Why do faculty rely on GRE scores?

Inherited practice & Lack of training in selection

Convenience

Mindsets & Evaluative scripts$^1$:

*Definition:* Stories that people tell themselves to justify taken for granted behavior

Faculty associate GRE scores and grades (conditional on curriculum rigor and institutional prestige of where the grades were earned) with intelligence, which they associate with Belonging in an elite intellectual community

*Risk* profile

1 Goffman, 1959; Lamont, 2009
GRE Scores & Intelligence

In interviews, 50% of the sample volunteered some idea about intelligence when asked what GRE scores signal (e.g., “sheer intellectual horsepower”, “native intelligence”)

In meetings, >50% of GRE mentions were what I classified as smart talk.

“Someone who does that well on the GRE is unlikely to be lame-brained. They are likely to be smart.” (philosophy)

“Freaking genius” (political science)

“I question she has what it takes.”

“[He was] from a different planet and we were confident that this person was not going to be one of us. He’s not going to be a full member of the scientific community.” (biology)
## Astrophysics committee

<table>
<thead>
<tr>
<th></th>
<th>Prabhat</th>
<th>Jeff</th>
<th>Juan</th>
<th>Wayne</th>
<th>Chris</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Assoc Prof</td>
<td>Assoc Prof</td>
<td>Assoc Prof</td>
<td>Asst Prof</td>
<td>Ph.D. candidate</td>
</tr>
<tr>
<td>Institutional affiliations</td>
<td>Ivy</td>
<td>Ivy</td>
<td>Ivy</td>
<td>Big Ten</td>
<td>Big Ten</td>
</tr>
<tr>
<td>Born</td>
<td>Int’l</td>
<td>Domestic</td>
<td>Int’l</td>
<td>Domestic</td>
<td>Domestic</td>
</tr>
</tbody>
</table>
Juan: Is it enough to be a woman in science?

(Discussion of how different perspectives might affect the community.)

Prabhat: Lisa said she wants to be a role model because she never received explicit encouragement until recently. She wrote about the importance of providing active support, not just the absence of discrimination.

Wayne: Shawna says she needs to develop self-confidence and overcome self-doubt.

Juan: And then there’s Amy, who claimed to experience teasing and bigotry from her peers and a high school science teacher. She went to an all-women’s college so she could still study science.

Chris: I’m less persuaded by that story. Maybe the teacher was young and inexperienced in handling high school boys. She might come to the program with an axe to grind.

Juan: Either way, now she’s taking action, organizing a lecture series on women in science... We need to read between the lines on these things.

(Person eventually nominated had started an astronomy outreach program and had letters of recommendation from familiar names.)
American higher education is racially polarized by institutional prestige.\(^2\)

1. Attiyeh & Attiyeh, 1997
2. Carnevale & Strohl, 2013
3. Posselt, et al., 2012
Inequality & Networks of Exchange in Higher Education

Danowitz-Sagaria, 2002
- Academics minimize perceived hiring risks by emphasizing “fit”
  - ...as a signal of conformity
  - ...as a mechanism by which shared networks affect the information deemed legitimate.
- “By privileging information from known sources, search committees effectively limit the range of information they were willing to consider” (p. 689).

Burris, 2004
- Academia as caste system.
- 70% of faculty in the top 94 sociology departments earned Ph.D.’s in the top-20 ranked programs.
- Networks of Ph.D. exchange...
  - ...generate discipline-specific social capital.
  - ...reinscribe prestige hierarchies.
Two views of trust

**In Rational Choice theory**

- **Def:** An expectation that others’ actions will be in my self-interest.\(^1\)
- Approaching graduate admissions as investment encourages risk aversion.\(^2\)
- Halo effect makes it easier to trust students from high status institutions\(^3\)

**In social capital theory**

- Value of social capital is in the structure of relationships among actors, including their trust.\(^4\)
- **Trust networks:** “Ramified interpersonal connections, consisting mainly of strong ties, within which people set valued, consequential, long term resources and enterprises at risk to the...mistakes & failures of others.”\(^5\)

---

1 Kohn 2008  
2 Klitgaard, 1985; Posselt, 2014  
3 Paxton & Bollen, 2003  
4 Coleman, 1988, p. 598  
5 Tilly, 2005, p. 12
Findings

Trust facilitates investment amid uncertainties.
1. “You have so little else to go on”: Uncertainty about academic preparation compels attention to prestige
Uncertainties about preparation motivate reliance on prestige.

- Ceiling effects in GPA reduce grades’ utility in distinguishing applicant preparation.
  - “Everybody in the pool has such high GPAs that it’s not meaningful.” (Astrophysics)
- Contextualizing GPA by perceptions of institutional quality effectively broadens the range.

![GPA Range Diagram]

3.5 at Unknown U.  
4.0 in Famous Dept.
What’s great about hiring professors is we have direct evidence of exactly what they did... It’s not easy, but it is information rich. Whereas I would say graduate admissions is information poor. So then one tends, or we tend, a lot to rely on signals that are low quality like, one of the frequently used ones is the quality or prestige of the undergraduate school. Lousy signal, I think.

Why is it valued so much do you think?

Because you have so little else to go on. You have grades, which I think are a good signal. But the people we admit are always going to be right around 4.0... What else do you have? So you have the tests, and yeah, we definitely sort of have an expectation of high scores on the test even though no one likes to use them. But increasingly, you have plenty of people who are really high on the test scores and really high on grades. Tons of those people. So now what do you use?

It sounds like you’re looking for variance.

Right, right. So you use the prestige of the school.
2. “But we just don’t know the letter writers”: Uncertainty about letters of recommendation and their authors
Ambiguity about hyperbole in letters.

• Letters help faculty fill in informational gaps about personality, research experiences/abilities, and socio-emotional skills

• But effective interpretation requires
  • ‘‘reading between the lines’’
    - noting both ‘‘what people say and what they don’t say’’
    - attending to tone.

Julie: What difference does it make to know the letter writer?

Bill: What it means is that we can really, truly evaluate their letters because you know them. (physics)
Praise was deemed more “trustworthy” and “reliable” when the letter was from a known source.

- Trust in a letter writer was often a bottom-line basis for decisions.
- In philosophy, it could be the final word:
  - “The letter was written by a graduate student. I don’t know how to read that.”
  - “I trust her [the letter writer] a lot, so it’s powerful.”
  - “But we just don’t know the letter writers.”

- What happens if a letter sounds like more than a professional endorsement?
Letters for international applicants

When it conformed to American norms, a letter could “speak to the record” or “compensate” for uncertainty about academic preparation. If it failed to fit the American style, faculty declared it “pretty useless” and “basically worthless.”

- senior professor, humanities

• “Now, 15-20% of our letters are from Asia. If you’ve ever read them, you can’t compare them with Americans. You basically have to ignore the letters.”
  - senior professor, humanities

• “Internationally, they [the committee] know the programs they trust, and there are some institutions producing very good economists, but their faculty don’t know how to craft the kind of recommendation that at US reader will find persuasive.”
  - admissions chair, economics
Blind spots in faculty assumptions.

- In both how letters are written and how readers interpret them.
- Some assumptions are gendered and racialized.
- Assumptions about risk are informal and subject to biases. For example,
  - Faculty place undue confidence in their own ability to predict who will be successful.
  - It’s difficult to reliably predict Ph.D. completion for populations who rarely enroll (i.e., problem of small N)
- Student outcomes result from what they bring to the table AND from the educational experience & climate we provide (Lovitts & Nelson, 2000). The same is true for faculty!
Biases in Letters of Recommendation

Trix and Psenka (2003) found that compared with letters written for men, letters written about women were:

• Shorter

• more likely to lack basic features, such as how they knew the applicant

• concrete references about the applicant’s record

• evaluative comments about the applicant’s traits or accomplishments.

• Less likely to be aligned with research record and ability.
Biases in Letters of Recommendation

Descriptive words that may be used differently in evaluating members of different social groups:

- **Grindstone words**: imply that women succeed more through effort than ability
  - *hardworking, conscientious, dependable, meticulous, thorough, diligent, dedicated, careful.*

- **Ability traits**: used less frequently for women
  - *talented, smart, able, capable, brilliant*

- **Communal adjectives**: negatively associated with hireability
  - *affectionate, nurturing, kind, warm*

- **Agentic adjectives**: used less frequently for women
  - *ambitious, dominant, self confident*

- **Standout adjectives**: repeated more for men than women
  - *excellent, superb, outstanding, unique, unparalleled*

Adapted from Leigh ADVANCE Best Practices for Reading and Writing Letters of Recommendation
Implicit bias
Milkman et al.: “What comes before”

• Field experiment compared faculty responses to email inquiries from prospective graduate students.
• Emails sent to 600 professors, identical in all ways except the name on the bottom.
• Professors responded significantly less often to prospective students whose names suggested they were Black, Latino, from Chinese, Indian, and/or female.
• And when they did respond, they took significantly longer.
• Effects strongest in private universities.
Strategies to Reduce Implicit Bias in Selection

- Devote adequate time.
- Avoid premature ranking (anchoring bias).
- Use a rubric or other evaluation form.
- Critically analyze supporting materials.
- Be transparent: What criteria? Are they the right criteria?
- Appoint diverse groups for file review and encourage maximum participation.
- Be accountable. Be prepared to explain your decisions.
Some current problems & needs

PROBLEMS
- Ineffectiveness in predicting success.
- Inequitably siphons out already underrepresented groups.
- Overly focused on achievements at the cost of potential.
- Unstructured review is opaque and enables biases to creep in.
- Inertia: faculty rarely receive support/incentives to evaluate and improve practices.

NEEDS
- Efficiency & utility at scale in addressing problems.
- Focused attention to diversity & non-cognitive factors.
- Transparency & accountability.
- Structures to enable holistic review.
- Social connectedness, competence, and autonomy to make change.
A Framework for Holistic Review

Comprehensive
Contextualized
Systematic
Three types of holistic review among undergraduate admissions staff:

*Whole file*: Considers all parts of the application.
*Whole person*: Considers many facets of the applicant.
*Whole context*: Considers the context in which the applicant achieved what they did.

Bastedo, Bowman, Glasener, & Kelly, 2018
Comprehensive

- Criteria should be numerous and diverse
- Consider
  - the whole person and their potential
  - diverse perspectives improve scholarly work
  - socio-emotional skills are necessary for outstanding professional performance
Contextualized

- **Metrics in context**: Note intrinsic errors; distributions of applicants’ opportunities
- **Achievements in context**: Interpret student achievements relative to societal patterns
- **Admissions in context**: How can applicants aid your program’s identity/mission or link to broader goals?
Systematic

- Center review on shared and predefined criteria with structured tools/protocols to enhance efficiency & consistency
- Build in opportunities for flexibility/nuanced views
- Build in safeguards & checks to promote equity and limit biases in the process
- Training and selection of gatekeepers
- Coordinate evaluation with recruitment/yield processes
We should concretely consider diversity contributions in graduate admissions & faculty hiring, given the practical benefits of diverse scientific teams.

Research Cited More
- Freeman & Huang, 2014

Better Problem Solving
- Phillips et al. 2008
- Page, 2007

Better Ideas
- De VaanStark & Vedres, 2011
- Burt, 2004