RETHINKING RESEARCH ASSESSMENT
UNINTENDED & COGNITIVE SYSTEM BIASES

DORA

Judgment and decision-making biases that impact how we weigh options and make choices have been shown to result in inequitable review, promotion, and hiring practices. **While recognizing these biases at a personal level is important, creating new structural and institutional conditions to reduce bias can be even more valuable.**

People tend to dismiss evidence that doesn't fit their initial judgments or preconceptions

Example: Cherry-picking information from a CV to confirm the view one already has, or dismissing potential warning signs because a candidate has already been accepted as a good fit.

Why it's problematic: Our initial conceptions are often based on subjective experiences and limited data. Failing to gather and consider counter-evidence makes us more likely to fall into old ways of thinking.

We often take the path of least resistance unless there are strong reasons not to

Example: Continuing to use citations from academic journals as a primary indicator of impact or quality, rather than considering alternate quantitative indicators of real-world value.

Why it's problematic: People often stick with recognizably flawed processes because the effort to fix them or adopt new ones is perceived as too much effort.

Resources often flow to those who already have them

Examples: Highly cited references may be more cited in part *because* researchers see that they're highly cited. Researchers with a long track record of grants receive a disproportionate amount of new funding.

Why it's problematic: When people lack the time or motivation to vet results, this can make access to resources even less equitable.

Once metrics are accepted as a way to gauge value, they start to lose meaning as objective measures

Example: Reward systems that rely on easily measurable qualities—like citations and publishing in high-JIF publications—can lead people to "game" the system.

Why it's problematic: When quantitative measures have an outsize impact on how people are rewarded, it can increase the temptation to focus on a narrow set of activities and reduce investment in other meaningful, but less rewarded, achievements.

Tackling these infrastructural and institutional implications of common biases can help promote and support more equitable practices:

"Objective" comparisons are not necessarily equitable

Qualities that can be measured or ranked are tempting because they feel less subjective, but can feed a false sense of precision.

What can institutions do?
- Balance the use of quantitative metrics with qualitative inputs, like narrative CVs, that capture more intangible qualities
- Select standards based on a wide set of inputs rather than a narrow or anecdotal set
- Recognize where setting specific, quantifiable goals may be reinforcing some behaviors at the expense of others

Incumbent processes and perceptions have the advantage

Many institutions have deep legacy traditions that become normalized over time, but these organizational habits can also keep new ideas and people out.

What can institutions do?
- Make the benefits of new behaviors concrete, salient, and easy to grasp
- Recognize where old assumptions may overly reward those who are more traditionally successful, at the expense of new or more diverse talent
- Set, publicize, and adhere to measurable goals that look beyond traditional norms of success when reviewing potential candidates to broaden the pool of individuals under consideration

Central diagram labels
- Matthew effect
- Status quo bias
- Campbell's law
- Confirmation bias
- Anchoring
- Availability
- Halo effect

We gauge value by association

Highly rated or prominent institutions and journals (and those associated with them) often get the benefit of the doubt based on familiarity or reputation rather than reality.

What can institutions do?
- Use structured interview protocols to keep decision-makers focused on agreed-upon qualities, rather than on reputation
- Explicitly articulate and consider long-term and qualitative values, as well as short-term or easily quantifiable needs
- Have applicants highlight and articulate their most meaningful contributions to reduce reviewer reliance on journal names or quantifiable characteristics of productivity

Individual data points can accidentally distract from the whole

It's hard to weigh all information equally, which can give initial or "shiny" data points and personal reference points an advantage.

What can institutions do?
- Assemble diverse teams—across gender, seniority, cultures, and under-represented minoritized populations—to bring a range of perspectives and experiences into decisions
- Look outside your institution or discipline to broaden a sense of "normal"
- Put reputation-based indicators like education at the end of applicant materials to reduce preconceived notions

Anecdotal, top-of-mind, or easily recalled data can inadvertently skew what we prioritize

Example: Prioritizing individual or memorable anecdotes, both pro or con, like getting a well known grant.

Why it's problematic: Overweighting information that more readily comes to mind may result in failing to consider other important evidence, missing the bigger picture.

We let positive impressions of individual attributes influence our overall opinions

Example: A candidate from a prestigious institution is thought to have more potential than one from a lesser known university

Why it's problematic: Giving preferential treatment to people based on inherited attributes may reinforce inequitable norms, which can lock out otherwise worthy candidates and fails to consider individuals equally.

The first piece of data we see or hear tends to set the bar against which we judge subsequent pieces of information

Example: Negatively comparing post-COVID-19 research productivity to pre-COVID-19, or using one's own personal life as a gauge to judge others' experiences

Why it's problematic: Initial anchor data defines the "normal" against which all other data is compared, which can skew our reference points by emphasizing relative comparisons between options rather than their actual value.