

# RETHINKING PEER REVIEW

**AVALANCHE OF APPLICATIONS** pushes federal agencies to make their selection processes less burdensome, more fair

ANDREA WIDENER, C&EN WASHINGTON

**HUNDREDS OF THOUSANDS** of grant applications have surged into the offices of federal funding agencies in the past few years. Ever-tightening government budgets mean fewer grants are given out, forcing scientists to write more proposals to sustain their laboratories.

That upswell has a major downside—and not just for anxious grant-seekers. Budget-strapped agencies must handle the growing landslide of applications with an unchanging amount of staff and time.

At the National Institutes of Health, almost 87,000 grant applications came in during the 2014 fiscal year. That's up from under 42,000 in 1998. And at the National Science Foundation, 48,000 proposals were reviewed in 2014, up 50% from almost 32,000 in 2001.

Closer to home for chemists, NSF's Mathematical & Physical Sciences Directorate is struggling with the deluge, says its

leader, NSF Assistant Director F. Fleming Crim. In 2000, the directorate got 4,500 proposals and funded 37% of them. In 2014, it received more than 8,000 proposals but could make awards to only 24%.

Agency officials say the increase in applications is on the verge of overloading the federal system. Not only are the program officers who shepherd the grants overwhelmed, but “with twice as many proposals, we have to get twice as many reviewers,” Crim says. “It is harder and harder for people to agree to review because they are getting so many requests.”

To cope, NSF and NIH are conducting experiments of their own to make peer review more manageable. At NSF, those trials include straightforward ideas such as holding more online, rather than face-to-face, review panel meetings and reducing the number of grants that panels discuss. But they are also trying more far-out schemes.

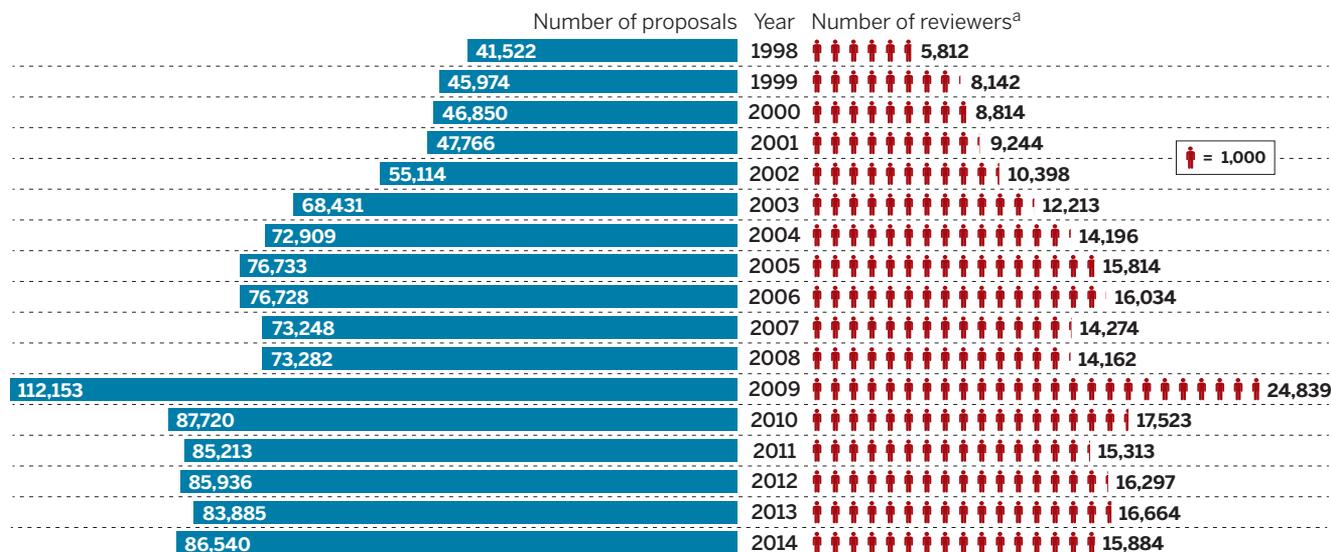
NIH is also trying new methods to lessen the burden placed on reviewers, many of whom serve multiyear terms on scientific review panels that meet primarily in person several times a year. In addition, the agency is working on pilot projects to make sure its peer review system is fair and that panels are choosing the best proposals.

Those pilots are designed to bring scientific rigor to bear on any changes in peer review, says Richard Nakamura, head of NIH's Center for Scientific Review (CSR), which manages 70% of the agency's applications—individual institutes handle the rest. Proving that reviewers actually choose the best applications for funding could help change the perception among many scientists that getting a grant is a crapshoot, he says. “As a review shop for a scientific organization, it seems crazy that we're not applying science to what we do.”

**“PEER REVIEW** is not very well described in the literature, which is surprising given that it is so fundamental to science,” says Stephen Gallo, who has studied peer review as technical operations manager for the American Institute of Biological Sciences. The nonprofit group organizes peer review panels for some federal agencies and nonprofits. “The only way that we are really going to make the system better and get the best research is to actually study it,” he says.

NSF and NIH are being cautious about

**GRANT GLUT** The number of applications NIH receives is reaching the limit of how many it can handle, says Center for Scientific Review Director Richard Nakamura. The number of reviewers NIH needs to handle the increase continues to climb.



**NOTE:** Bump in 2009 data attributable to onetime funding from the American Recovery & Reinvestment Act. **a** CSR reviewers only. **SOURCE:** CSR

how they conduct the pilot studies because they know researchers' funding is on the line. They are trying out small studies spread over several years. "We want to make sure we don't damage the integrity of the process," explains NSF's Stephen Meacham, who co-led a working group that examined the agency's peer review system.

**VIRTUAL PANELS** are something that both agencies have been trying for several years. These panel meetings allow for more flexible scheduling and can cost less than in-person panels.

Lawrence R. Sita, a chemistry professor at the University of Maryland, prefers them to in-person panels—he likes that he can talk science while sitting at his kitchen table. But in any case, Sita says a meeting where scientists can discuss applications with each other is important. "You see who has done their homework and who hasn't," he says.

At NSF, virtual panels have increased the diversity of panel members by allowing those with family responsibilities—especially women—to participate, Meacham adds.

NSF's working group explored a number of ways they could change peer review before settling on seven pilot efforts to try out. Virtual panels were one option. Another invited scaled-back preliminary proposals that included just the science portion of the application. That way, scientists whose proposals are screened out wouldn't have to complete the tedious work of presenting their methodology or budget.

A pilot that proved popular involved using chat rooms or virtual communities such as Second Life to allow some discussion about grants before reviewers participated in a panel meeting. They conducted much of the discussions about applications online, so when they got to the actual panel meeting, "they could focus on the ones that were in the difficult middle," Crim explains.

Another way to scale back peer review work involves reducing the overall number of applications. One pilot project involved a switch from a single deadline to rolling acceptance of applications. The number of applications fell from 173 to 85.

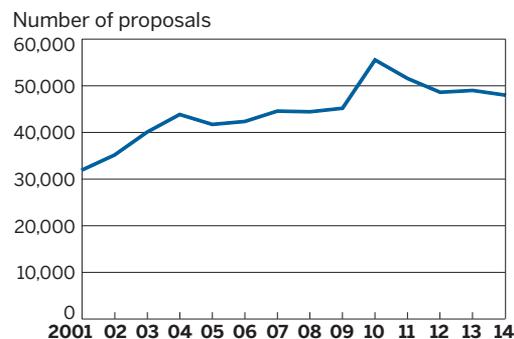
One of the most provocative pilots involved something that is normally taboo in grant review: recruiting reviewers who are applicants for the grant. George A. Hazelrigg, deputy director of the NSF Civil, Mechanical & Manufacturing Innovation Division, ran this pilot panel, which as-

signed applicants to review proposals from seven of their fellow applicants.

Hazelrigg and his colleagues used game theory to create rules to prevent people from downgrading their competitors. "It is designed to have people do the review, but if they aren't honest, it hurts them," he says. Of the 131 researchers who participated in this effort, only one dropped out.

"That is an interesting way of placing the burden of review on those who are applying," remarks Pennsylvania State Universi-

**PROPOSAL PROBLEM** At NSF, the number of grant applications received has increased by 50% since 2001.



**NOTE:** Bump in 2010 data attributable to onetime funding from the American Recovery & Reinvestment Act. **SOURCE:** NSF

ty chemistry professor Thomas E. Mallouk.

At NIH, many of the pilot projects started after Nakamura took over CSR in 2012, so the results aren't available yet. Although the number of reviews is a huge challenge, Nakamura's top priority when he became director was to examine diversity. Specifically, he was charged with finding out why African Americans are less likely to get funded than other racial or ethnic groups.

Nakamura and his colleagues are analyzing the text of grant applications to see if any wording might lead to African American applicants being screened out. They have also designed a pilot study aimed at eliminating bias by anonymizing applications as much as possible. This would cut major pieces of these documents, such as the reference list, biosketch, and applicant's institution.

"I like the idea of making the application blind. There is a huge amount of unconscious bias that exists," says chemistry professor and regular NIH reviewer Steven C. Zimmerman of the University of Illinois, Urbana-Champaign. He admits it would be difficult to make the selection completely blind, though.

Other pilots are trying to ensure that the agency is funding the best science, Nakamura explains. One project is addressing grade compression—the clustering of scores for applications around the top rankings, rather than panels spreading the grades out over the full range so the distinction between grants is more obvious. In addition to their normal grades, the project is asking reviewers to rank their top 10 favorite grants. Another pilot involves a second review of some grants to see if the scoring system consistently leads to the same results.

Making choices among grants is especially difficult given the decline in funding levels, Penn State's Mallouk says. When grants are awarded to fewer than 20% of the applicants, "the whole system goes haywire. It's no longer really possible to differentiate the top 15 to 20%," he says. "There is a lot of chance and nit-picking in the system."

Mallouk suggests that agencies limit the number of projects they will fund on any particular problem, keeping reviewers from jumping on the bandwagon of whatever topic is suddenly popular. NIH isn't doing that right now, but it is attempting to judge whether the system of assigning grants evenly across its 173 topical study sections results in awarding the best grants.

"Are our best applications unevenly distributed among study sections?" Nakamura asks. "The answer is probably yes."

All of these attempts to examine the review process will be difficult to implement if the system continues to be overloaded. Nakamura would like to reduce the annual number of applications to 60,000, or even 40,000. NIH is exploring ways to do that, he says, including possibly limiting the number of grants any one investigator could receive.

"I wouldn't object to having some limit to the number of research grants," Zimmerman says. It ultimately might mean that some academic research groups would be smaller, which he points out would leave more time for mentoring.

For NIH, the application number problem grew worse in 2014, when the agency began to allow scientists to revise and resubmit applications. That policy change will be fully implemented in 2015, and it will undoubtedly bring in more applications.

"The question is, Will that recede, or will that continue?" Nakamura asks. If the policy continues to increase the number of applications, "then we're in real trouble." ■