Annals of the Entomological Society of America, 110(5), 2017, 484–491
doi: 10.1093/aesa/sax059
Forum

Special Collection: Science Communication

Moving Beyond Metrics: A Primer for Hiring and Promoting a Diverse Workforce in Entomology and Other Natural Sciences

Robert L. Reece[1] and Margaret C. Hardy[2,3]

[1]The University of Texas at Austin, Department of Sociology 305 E 23rd St, Austin, TX 78712 USA (Robert.L.Reece@gmail.com; Twitter: @PhuzzieSlippers), [2]Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, 306 Carmody Road, Brisbane 4072, Australia (Twitter: @DrMaggieHardy), and [3]Corresponding author, e-mail: nrl4@cdc.gov

Subject Editor: Lawrence Hurd

Received 20 May 2017; Editorial decision 23 June 2017

Abstract

The lack of a representative scientific workforce is a challenge being addressed by federal agencies, funding bodies, and research institutes. As organizational units emphasize the work of training, recruiting, retaining, and promoting an inclusive scientific body, issues around bias and metrics appear frequently in the literature. The goal of this manuscript is to introduce some of the most commonly misunderstood areas of bias, and ways to adjust hiring and promotion practices; and to focus on innovative ways to quantify scientific outputs through metrics other than journal impact factors and paper citation numbers. These alternative metrics, or altmetrics, are increasingly reliable predictors of impact for a scientific community that is asked to engage more frequently and more effectively with the public. Additionally, we address discipline-specific concerns around metrics used for hiring and promotion in biosecurity and biosafety, taxonomy, and researchers in government and regulatory roles. We conclude that in the future, societal and scholarly impact will be more closely tied. This is a critical time to ensure scientists are adequately trained and recognized for their contributions as public intellectuals, and supported in this important work.

Key words: affirmative action, career development, equity, diversity, STEM (science, technology, engineering, and mathematics)

When applied to the hiring and promotion of research scientists, metrics have many of the same problems as GDP (gross domestic product), because both rely on quantifying productivity (Abramo and D'Angelo 2014). GDP is an economic concept of value added, calculated based on the market value of recognized goods and services over a period of time, and is often used incorrectly as a proxy for determining a country's standard of living (Giannetti et al. 2015). GDP leaves out critical volunteer and unpaid domestic services (like child rearing), as well as quality of life measures like leisure activities and a clean environment (Anonymous 2016). It is one measure of the material standard of living, but is not the most important measure that indicates the productivity of a country's economy. This is why we use a constellation of different economic measures to assess progress over time against a number of goals that improve the quality of life in a country or region (Costanza et al. 2016).

In academia, teaching, research, and service are weighted against the percentages assigned in a position description; in industry or government jobs, research or public engagement might be more important, particularly with policy or extension roles (Oakley and Garforth 1985). Although we also have an array of metrics used to demonstrate productivity, there is a gap in our understanding of how the metrics are related to one another. As we discuss in this manuscript, the bias in how the metrics are applied and constructed has a disproportionately negative impact on researchers who are traditionally underrepresented in the natural sciences. Individuals who wish to identify their own areas of bias can consider taking the free Harvard Implicit Bias Test (https://implicit.harvard.edu). If our goal is to advance the discipline as rapidly as our global challenges require, we need a sustainable, diverse workforce trained and retained.

Although the number of underrepresented minority (URM) graduate students earning Doctoral degrees is increasing, the demographics of employed scientists illustrates the depth of challenges faced by today's emerging leaders (Fig. 1). In the United States, although URM graduate students earn Doctoral degrees at a rate that is roughly representative for their race (according to US Census data) they are employed at a lower rate; conversely, white Americans earn fewer PhDs than their census population percent, but are employed at rates near or above their census population

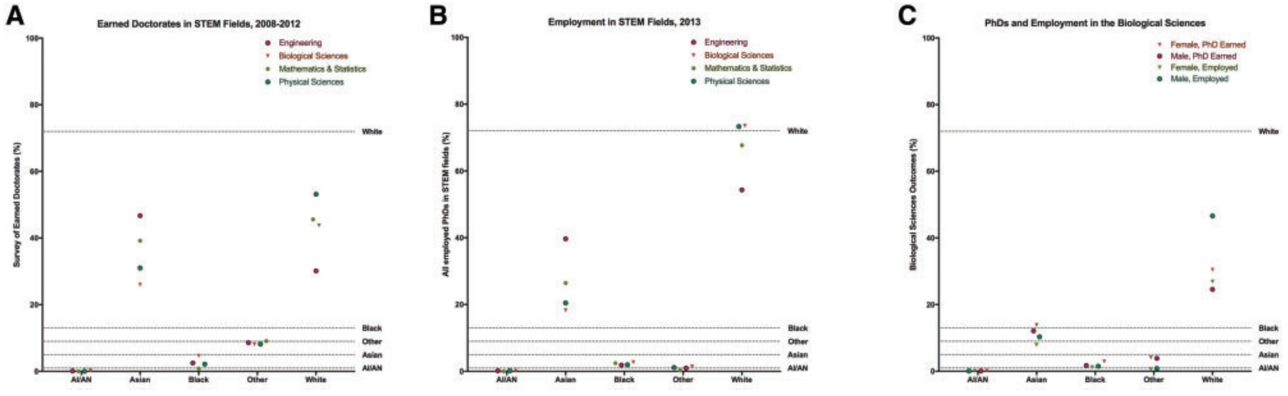

Fig. 1. A comparison of PhDs earned in STEM fields (**A**) and employment in STEM fields (**B**), with a specific comparison of PhDs and employment in the biological sciences (**C**) in the United States. Census data categorizes respondents as "male" or "female," and by standardized race and ethnicity categories. Native Hawaiian and Pacific Islander persons are included in "Other," as are persons who identify as multiracial. AI/AN—American Indian or Alaska Native. The dashed line represents the reported population percentage in the 2010 US Census (data courtesy the US Census Bureau), which was the most recent year the data were collected. The employment and earned doctorate data were sourced from: NSF/NIH/USED/USDA/NEH/NASA/NCSES and are accessible at: <https://ncses.norc.org>. The complete data set is available on Figshare: <https://doi.org/10.6084/m9.figshare.5027447.v1>.

percent (Fig. 1A, B). These challenges are magnified for women in STEM across disciplines, and amplified for transgender and nonbinary persons (Falk and Long 2016, Wang 2016), although most surveys neglect to collect data on gender identity other than "male" or "female." For the biological sciences alone, females earn Doctoral degrees at roughly the same rate as their male counterparts. Further, white men are far more likely than white women, or men or women of other races, to be employed (Fig. 1C). Although efforts to make the PhD system more inclusive have been somewhat successful, even though qualified candidates from diverse backgrounds are being produced employment demographics do not proportionally reflect this diversity. There are a number of interconnected reasons for these outcomes, but the focus of this paper is to explore how STEM can be more inclusive in its hiring and promotion practices.

The goal of this manuscript is to outline the benefits and limitations of commonly used metrics. Specifically, we explore ways that bias reduces our ability to be inclusive and to build a representative scientific workforce. Because altmetrics help measure outputs that include mentoring and outreach programs, our hypothesis is that a more representative cohort of scientists can be hired and promoted. By arguing that these measures should be interpreted as largely accurate measures of different output types from a single researcher, we explore other ways to measure impact. Additionally, we aim to highlight how these metrics can support the training, hiring, retention, and promotion of outstanding scientists who are engaged in essential capacity-building initiatives.

Manufactured Metrics: Strengths and Shortcomings

Like members of other professional fields, scientists are under increasing pressure to quantify their achievements according to external standards. However, as with other metrics discussed in this manuscript, publication output metrics are biased against women researchers (Symonds et al. 2006, Ghiasi et al. 2015). A meta-analysis of eight million manuscripts across a range of natural and social sciences revealed that men are more frequently found in the high-status first and last author positions, and women are significantly underrepresented on single-author papers (West et al. 2013). This holds true in the commercial space as well: although more women researchers are filing patents now than in the past, women faculty members patent at about 40% the rate of men in the same roles (Ding et al. 2006).

To enable researchers and their employers to quantify the outcomes of taxpayer funds expended on research and development, a number of companies have developed subscription services to provide media monitoring and provide yet another proprietary measure of impact (Carpenter et al. 2016). Although interdisciplinary research is encouraged, there are few mechanisms in place for the standardization of discipline-level metrics, including the widely used Thomson Reuters Journal Impact Factor (JIF). The JIF is of limited use when comparing journals that are not standardized with respect to discipline and scope (Seglen 1997); further, considerable differences in JIF have been reported between journals that primarily publish review articles, and those which publish original research (Lariviere et al. 2016). In 2015 the highest impact factor journal in entomology earned a score of 13.5 (*Annual Review of Entomology*), but is a journal that publishes review papers only; the highest ranked journal publishing original scientific research had an impact factor of 3.8 (*Insect Biochemistry and Molecular Biology*). In the field of cell biology, for comparison, the highest impact factor journal was ranked 38.6 (*Nature Reviews Molecular Cell Biology*); the top ranked journal publishing original research received an impact factor of 30.4 (*Nature Medicine*). As funding bodies and administrators increasingly rely on metrics like impact factor to assess the quality of research contributions, being able to publish in high-impact factor journals has considerable influence over a researcher's ability to be competitive for funding and promotion opportunities. This is particularly true for interdisciplinary funding applications, where each researcher's track record must contribute to the success of the project; scientists representing smaller disciplines or fields with fewer high impact factor journals can be at a disadvantage when their track record is compared to those of researchers in larger, higher profile research areas.

Research Councils UK (RCUK) defines research impact as "the demonstrable contribution that excellent research makes to society and the economy," and can include academic, economic, and societal impact (for more information on how RCUK defines and assesses research impact, see http://www.rcuk.ac.uk/innovation/policies/). In Australia, the 2015 National Innovation and Science Agenda (innovation.gov.au) called for a national impact and engagement assessment to be introduced; 2017 is the pilot year for the project. Federal research agencies in the USA require a justification of the intellectual merit of a project, as well as a description of the broader impacts (the potential of a proposed project to benefit society and contribute to the achievement of specific, desired societal outcomes). One of the key challenges is to ensure the broader impacts of a project are suitable and appropriate for the communities that are involved. Researchers should provide programs that meet the community need rather than fulfilling the grant proposal guidelines without first considering the need of the community partners, and programs should include assessment and evaluation as a central component of the work.

An overview of the intersections of societal and scholarly impact, along with several commonly used measures of each, helps illustrate the ways societal and scholarly impact can be integrated (Fig. 2). As research scientists try to navigate the often-disparate metrics that are used in hiring and promotion, scholarly impact is often seen as in conflict with societal impact. Each of the sectors has a role to play in creating a research narrative that an individual can use to explain their impact, and when committees review candidates the person description should consider which elements of research impact are most important for the role. Further, it is worth noting that the diagram focuses on research impact, and does not consider other essential work like teaching and mentoring.

Fig. 2. Research impact comes in many forms. Ideally, scientists try to maximize the societal and scholarly impact of their research, but the emphasis on impact factor and citation metrics for promotion and hiring can reduce a scientist's bandwidth to engage in activities with a high societal impact. The use of altmetrics and creating engagement and outreach programs with appropriate evaluation and assessment in place can help ensure these activities are assessed as central to the work of research scientists.

Challenges With Traditional Metrics: Recognizing Bias

Gender bias has been identified in hiring across a number of research disciplines. Men are twice as likely as women to be hired for a mathematical task, based on appearance alone (Reuben et al. 2014). High-achieving male faculty in the life sciences train 10–40% fewer women in their laboratories compared to other investigators (Sheltzer and Smith 2014). Even at the undergraduate level, faculty are more likely to consider a male undergraduate science student as significantly more hirable and competent for the role of a laboratory manager position than a female undergraduate science student (Moss-Racusin et al. 2012). Taken together, these studies indicate gender bias begins early in the career-building process and continues through the critical trainee and postdoctoral phases.

At the elite level, the visibility of diverse role models is correlated with a diverse student body (Birdsall et al. 2017). Despite playing a considerable role in the hiring and promotion of instructors, the role of gender bias in student evaluations of teaching (SET) has been ambiguous until recently. In a study of an online anthropology/sociology class in North Carolina, students rated the male teacher identity significantly higher than the female, regardless of the instructor's actual gender identity (MacNell et al. 2015). A large multinational study of first-year university subjects showed SET results rely more on the gender bias and grade expectations of students than they do on teaching effectiveness: according to the authors, "SET are biased against female instructors by an amount that is large and statistically significant… It is not possible to adjust for the bias, because it depends on so many factors" (Stark and Freishtat 2014, Boring et al. 2016). Male undergraduates have also been shown to underestimate the performance of their female peers in an biology course (Grunspan et al. 2016), which has implications for grading in courses that rely on peer assessment.

After strong lectures, students ranked gay and lesbian university instructors more negatively, perhaps making it a challenge for high-performing gay and lesbian instructors to be recognized for teaching effectiveness (Ewing et al. 2003). In other studies students perceive a gay teacher as significantly less credible than a straight teacher (Russ et al. 2002). Further, homophobic harassment in the workplace has been shown to be widespread against instructors or academics that identify as lesbian, gay, and/or transgender (Irwin 2002). Within the past year, >35% of LGBT physicists considered leaving their department; 50% of transgender and 22% of cisgender respondents in the same survey reported exclusionary or harassing behavior in the workplace that could interfere with career advancement (Falk and Long 2016).

The Impact of Policy Shortcomings on Personal Lives

Nowhere are the politics of a research field felt more keenly than when there is sexual violence. A recent survey of researchers showed that at field stations both men and women respondents experienced harassment, in the form of inappropriate comments (71% of women respondents compared to 41% of men) and assault, in the form of unwanted physical contact (26% of women respondents compared to 6% of men), with trainees of both genders at a greater risk than established scientists; fewer than 40% of respondents worked at field sites with a code of conduct (Clancy et al. 2014). An ethical code of conduct can be established and implemented for professional societies as well as for department-level communities (Hardy 2014, 2016), in concert with the legal and organizational protections offered by an institution. In 2015, the nonprofit collective EntoAllies (https://entoallies.org/) was formed to prevent harassment and bullying at entomology conferences. The group remains an excellent resource to support all those in the entomology community.

Such organizations help maintain safe academic spaces, which are invaluable in facilitating quality work. When researchers feel unsafe, their work is inevitably compromised. And this safety must extend farther than the explicit violence of sexual harassment and assault to more subtle slights and microaggressions (Constantine et al. 2008). Organizations offer validation in spaces where it may be difficult to prove the existence of bias and microaggressions. And the struggle for proof is another indirect way that feeling unsafe stunts research progress and quality. In his classic 1970 essay, "The Furious Passage of the Black Graduate Student," Douglas Davidson discusses the barriers black graduate students face as they attempt to complete their graduate studies (Davidson 1970). He notes:

> "The difficulty of 'proving' that an individual is racist reflect the insidious nature of 'civility' and 'tact' in the American way. It is his awareness of and sensitivity of the operation of covert, institutionalized racism that accounts for the 'furious passage' of the Black student through graduate school" (p. 194).

This furious passage, which he describes as "extreme inner turmoil," does just as much to diminish the research capability of unrepresented scholars as the oppression itself. The next section will discuss the importance and ways of creating an inclusive, diverse environment.

Defining Diversity and Inclusion

In the United States in 2013 (the most recent numbers available on this), 78% of full-time faculty were white. Already out of proportion with their proportion of the country's population, those numbers became progressively more lopsided as the number of persons obtaining the rank of professor increased. Of full professors, 84% were white and 69% were men, relative to 74% and 48% respectively at the level of assistant professor (National Center for Education Statistics 2016). This jarring disproportionality immediately generates calls to increase the number of underrepresented minorities in full-time faculty positions. Particularly, it is important to increase the numbers beyond shallow tokenism, which centers on the belief that having one minority in each department is "enough" (Niemann 1999). However, academia must also acknowledge that diversity, even representative diversity, is more important than mere numbers. Indeed, rapidly increasing numbers without addressing the fundamental institutional structures and cultures that lead to such underrepresentation may generate a backlash from the groups that have typically dominated numerically (Yoder 1991). Therefore, if institutions plan to offer an environment for underrepresented faculty to succeed, it is vital that a focus on diversity and representation run parallel to a focus on institutional change and inclusion.

Developing an inclusive atmosphere requires acknowledging that underrepresented faculty often bring different knowledge and motivations to academia that do not always conform to traditional ways of being an academic. It is important to strike a balance between accommodating those differences and offering support to facilitate the transition to traditional norms. For example, in a review of two decades of research on faculty of color, the authors write that faculty of color are more likely to enter academia because they love teaching and serving their communities, including students (Turner et al. 2008). In an attempt to acknowledge and value this unique aspect of underrepresented faculty, the University of California system instituted a policy that weighs applicants' "contribution to diversity" as qualifications in the hiring process. However, underrepresented faculty often report being exclusively saddled with a bevy of diversity service work, including committees and mentoring students, which can be particularly stressful if a faculty member is the only, or one of few, URM in the department.

At this point, the goals of representative diversity and inclusion intersect. Underrepresented faculty feel more included when institutions demonstrate a sincere effort to achieve representative diversity, of both students and faculty, through institution-specific recruiting and programs that enrich the pipeline thus increasing the availability of applicants across the discipline (Turner et al. 2008). According to recent research, "across disciplines... faculty hiring follows a common and steeply hierarchical structure that reflects profound social inequality" (Clauset et al. 2015). The "leaky pipeline" that leeches women from careers in STEM has been shown to happen before the undergraduate degree commences and in the post-PhD job market (Miller and Wai 2015), and rewards that stimulate field commitment (including advanced training and high levels of job satisfaction) are lacking for women who pursue careers in STEM (Glass et al. 2013).

Ultimately, as the number of underrepresented faculty increases, they experience less isolation and more inclusion, feel less like a token, and are able to more evenly distribute diversity-related service work (Niemann 1999, Turner et al. 2008). With that in mind, hiring underrepresented faculty is only a first step in maintaining representative diversity. Faculty retention is the next step, and we discuss that later in the paper.

However, working toward an inclusive academic environment is fraught with pitfalls, even among seemingly well-meaning groups. In a recent work, researchers chronicled the challenges faced by a campus organization working to rectify the racism experienced by people of color at the university (Rodriguez and Freeman 2016). While the organization yielded some success, and was received relatively positively by campus administration, they report the difficulties of maintaining the focus on race and racism throughout the process as increasing numbers of white people joined the organization. The organization met two primary barriers: the first was a general reluctance among its members and the university to acknowledge that racialized outcomes, such as the underrepresentation of people of color on campus, were the result of racialized processes and racism; the second related barrier was white people's calls for "intersectionality" as a way to deflect from white culpability in racist events and to, indeed, minimize the role of race altogether. According to the authors,

> "...discussions of racism and white privilege were objected to because they did not allow white subjects to highlight their own experiences of inequality...intersectionality operated as a tactic to undermine the specific experiences of students of color by incorporating them into a framework in which diversity is defined in such a way that everyone is positioned as marginalized and subordinated" (p. 77).

While intersectionality is an invaluable concept that seeks to acknowledge the complex identities and social locations of people, it can be counterproductive when wielded like a bludgeon to minimize the complaints of certain groups.

Instead, an intersectional approach to diversity should seek to understand the multiple and compounding types of representation necessary to create a truly diverse and representational environment. For example, while women represent 31% of full professors in the United States, women of color comprise <4%. In this case, neglecting to examine intersecting identities would misrepresent the diversity of the workforce.

Measuring Phenomena: Altmetrics in Practice

Altmetrics are conceptually alternative (to citation) metrics (Haustein et al. 2016). Aggregators include Altmetric.com, Plum Analytics (which tracks output "artifacts," and is now part of Elsevier), and ImpactStory. Altmetrics can be easily integrated with any story about research or researcher impact, and provide quantitative feedback that can be useful for those in the position to hire or promote researchers who need a widely available standard to stratify applicants along an axis of engagement.

Recent scholarship has illustrated the depth of the public impact of social media. Tweets within the first three days of an article's publication (the "twimpact") have been shown to correctly predict articles that are then highly cited in the scholarly literature for digital medical (Eysenbach 2011) and ecological research (Peoples et al. 2016). Interestingly, the latter paper showed that social media can help articles from lower-impact factor journals become highly cited, but a researcher's strong social media presence alone does not make a paper highly cited. Scientists are highly trusted by the public, with 84% and 76% of Americans surveyed in a Pew Research Center poll expressing "a great deal" or "a fair amount" of confidence in medical scientists and scientists respectively to act in the public interest (Funk and Kennedy 2016). The medical community has recognized the importance of using social media to engage patients (Schnitzler et al. 2016) and particularly to increase parental awareness around certain conditions (O'Kelly et al. 2017). Exposure to misinformation on Twitter was shown to be correlated with reduced vaccine coverage (Dunn et al. 2017), further emphasizing the value of scientists engaging with the public on social media to disseminate accurate information.

In addition to indicating impact factor or the number of citations a paper has received, early- and mid-career researchers could include details for papers with a high altmetric score and description of wider media attention the paper created. The relevant media and communications department for an organizational unit generally has a subscription for a media monitoring portal that should give an indication of media attention that was created from an article, and librarians are gatekeepers for a vast amount of knowledge on how to measure and monitor altmetrics. With this expertise, altmetrics can be used to create a narrative about an individual researcher or specific project outcomes.

One of the obvious limitations of relying on altmetrics is that it relies on how much people talk about your work, in the scholarly literature (through citations), on social media, or through traditional news outlets. In science, unfortunately, it is not possible to say "any press is good press"—and altmetrics and social media analytics don't widely have the technology to distinguish whether the coverage is positive or negative, or even accurate. There are a number of platforms that can help a researcher follow the dialog, and potentially offer opportunities to correct the record or engage more broadly with stakeholders. There is an Altmetric.com bookmarklet that allows you to see the altmetric score for any scientific article; most journals and repositories now also report the altmetric score along with traditional metrics. ImpactStory is a free tool that enables tracking of researcher "achievements" linked to their ORCID and social media accounts, including the percentage of articles that have been published open access; the global reach of a researcher's work; and other online mentions that might be useful in creating a resume or promotion package.

Altmetrics are particularly useful as an "early indicator" of high-impact outputs, which is particularly important for early- and mid-career researchers (which may include research higher degree students, postdoctoral researchers, and nontenured faculty of equivalents). Specific recommendations for early-stage researchers include the use of a platform to record contributions to peer review (options include Publons or AcademicKarma); using a preprint repository, noting which discipline-specific repository is most suitable (options include bioRxiv and SocArXiv); and exploring the various options available to support open access publication (options include SPARC and journal-specific schemes, as well as those from philanthropists and federal funding schemes). A recent survey of Australian early- and mid-career researchers showed the need for training and continuing education in public engagement with science and effective social media use, illustrating the importance of creating training and praxis options in this space (Hardy et al. 2016).

Altmetrics can be particularly useful for researchers who have career interruptions or primary caregiving responsibilities. One of the clear advantages is the ability to use different social media platforms depending on an individual's career goals and research interests. One of the authors (MCH) is the mother of three children under the age of five, including a set of twins, and found social media to be greatly useful during her two periods of maternity leave. Social media allows time-poor researchers to stay up-to-date with current developments in their field, and every leading journal has a Twitter feed that provides 140-character summaries of the most recent published articles. Further, by having an open dialog on social media about a scientist's research, it allows the public the opportunity to engage and also raises the profile of the project.

Special Considerations

Although most job applications do not ask for citation numbers or impact factors explicitly, it would be reasonable for candidates to include a discussion of one or more scholarly outputs that have had an impact on the public, with support from altmetrics and other indicators (online readers, or other media coverage) as appropriate. These considerations should be taken as part of a larger conversation about how much "service" matters in hiring and promotion, and who currently does the bulk of this work (Misra et al. 2011, Misra and Lundquist 2015).

An overview of the general areas where entomologists and natural scientists might have special considerations (Table 1) highlights several themes. Public engagement and infrastructure that is relevant to the discipline, particularly in areas of biosecurity concern like taxonomy and biosafety, are often an ambiguous aspect of hiring and promotion processes.

Although not an exhaustive list, these are some of the topics that scientists may need to address during the hiring and promotion process. Determining ways to measure and report on the impact of these essential contributions is an ongoing dialog in the scientific community, but one that warrants further consideration to ensure equitable outcomes for individuals taking on these additional responsibilities.

Evidence-Based Suggestions to Recruit and Retain Diverse Scientists

In 2010, American Historically Black Colleges and Universities (HBCUs) produced >35% of all Black undergraduates who earned degrees in the physical sciences, and 19% of the 9% of overall Bachelors degrees in STEM awarded to Black students; the same year, HBCUs produced 10% of all Black STEM doctoral degree recipients (Upton and Tanenbaum 2014). Providing additional resources to institutions that are already working to make higher education and careers in STEM more inclusive could be a highly effective way to diversify STEM leadership.

The language used in job advertisements has been shown to be a key factor in successfully recruiting diverse candidates for STEM jobs (OConnell and Holmes 2015, Krome 2016, Wille and Derous 2017). Other manuscripts have provided comprehensive discussions of techniques and language that can be used to write effective job advertisements (Fine et al. 2014, Clauset et al. 2015, OConnell and

Table 1. An overview of specific areas of expertise from entomologists and natural scientists that are often challenging to assess in hiring and promotion processes

Area of expertise	Examples of outputs
Biosafety	Ensuring compliance with institutional and federal ethics policies as well as quality management and assurance. Data management and curation. Laboratory safety and participation on relevant institutional committees.
Commercialization and intellectual property management	Patenting, obtaining competitive funding, and grantwriting. Also covers entrepreneurship and collaboration with industry, government, and other stakeholders.
Extension and outreach	Stakeholder engagement and relationship management, particularly through programs targeted to a specific demographic (e.g., for students or growers). Structured evaluation and independent assessment of outcomes.
Government/ Regulatory roles	Some potential overlap with safety/biosafety, but quality assurance, diagnostics, and biosecurity. Compliance with legal obligations to federal legislation like CLIA (Clinical Laboratory Improvement Amendments).
Modelling and software development	Particularly when added to an open access repository, like GitHub or figshare, with permalink and source code.
Public scholarship	Social media contributions, public dialog, and contributions to traditional media outlets. Metrics from individual social media platforms can be included where informative. Featured work in aggregate sites that pull together collections of tweets or social media outputs on a topic and publish them as an article.
Resource infrastructure	Creating and maintaining infrastructure that advances the discipline, like databases, software, listserves. Includes development and maintenance.
Safety	Managing and processing safety violations (including sexual assault and harassment, as well as bullying or other forms of discrimination). Serving as an advocate or ombudsperson for students and trainees.
Taxonomy	Monographs, ad hoc species identification, and other relevant contributions that may not be highly cited or easily quantified. Expertise with particular species or characters, and management of museum collections or online datasets. Metrics might include number of identifications or requests.

Holmes 2015, Smith et al. 2015), and we encourage the reader to peruse the available resources within institutional structures like human resources. There are also a number of discipline-specific groups to support professional development for URM scientists, both formal and informal, that can help advertise positions widely. A number of long-standing mentoring programs could be modeled or brought on site, like the University of Oregon-based COACh program (https://coach.uoregon.edu/), to support women in STEM, and the Na Pua No'eau program in the Hawaiian Islands (Gentz et al. 2005). Mentoring programs that are structured and regularly evaluated provide opportunities for URM who conduct the bulk of this type of valuable activity to be compensated for their work, and the program inputs and outcomes can provide useful information during hiring and promotion reviews.

Despite the proven efficacy of mentoring programs to recruit and retain diverse scientists, service and mentoring in particular falls disproportionately on those already underrepresented in STEM. At the faculty level, the burden of service falls disproportionately to women, and with mentoring on women of color specifically (Misra et al. 2010, 2011, Misra and Lundquist 2015). These data, from a survey conducted by The Massachusetts Society of Professors from 2008–2010, showed that associate professors spend their weekly service hours on service to the university, profession, and on mentoring in different ratios depending on race and gender. The data showed white women (with a university:profession:mentoring service ratio of 12:5:11) and women of color (university:profession:-mentoring service ratio of 12:6:14) differ mostly in the amount of time spent on mentoring, while white men (service ratio 5:6:8) spend far less time on service than women or men of color (service ratio 9:8:9). In many disciplines, particularly the physical sciences, there are fewer women faculty and even fewer faculty that are women of color, causing this trend to have a considerable impact on hiring, retention, and promotion.

Institutional membership in one or more inclusive organizations provides benefits to individual researchers as well as to the organizational unit (for example, for employment advertising purposes). There are many excellent local, national, and international groups, but United States-based entities may want to consider the Society for Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS, https://sacnas.org/) or the American Association of Blacks in Higher Education (AABHE, http://www.blacksinhighered.org/).

Mentoring is another way to ensure faculty that are recruited are adequately supported in their professional development. This is a low or no cost activity, and can be implemented readily across disciplines. A mentoring program should take care to have some structure around responsibilities and expectations, but can take the form of a group or individual meeting on a periodic basis. Literature has shown that URM faculty at predominantly white institutions (PWI) are particularly well-supported by mentoring programs with three themes: 1) accumulating social capital and professional networks; 2) increasing the perceived value of faculty research areas and public scholarship; and 3) connecting with mentors who understand the specific concerns of URM faculty (Zambrana et al. 2015). Systemic, institutional development programs centered on faculty engagement, empowerment, and reward have also been shown to intentionally develop diverse scholarly communities in STEM (Whittaker and Montgomery 2014). Mentoring programs can ensure that staff apply for promotion and, where appropriate, tenure as soon as they are able and prepared to do so. These programs can also support the traditional metrics of scientific success, by providing internal grant reading panels and support for funding open access or high-impact publications.

In many instances, a pay audit can help ensure equity across salary scales. Most American universities already collect data on the pay differences between men and women faculty, and it would be relatively straightforward to annually review the salary of professional and faculty staff to be sure pay is unbiased. A research-intensive Australian university found a gendered pay gap favoring men of 15% for academics and 12% for professionals (Currie and Hill 2013). This process need not inherently interfere with the tenure process, as the salary across a pay band can fluctuate based on expertise and outputs.

What's Working Now

The Athena SWAN Charter is an initiative in the United Kingdom designed to increase gender equity for employment in higher education and research using a tiered award system of Bronze, Silver, and Gold. The Charter is designed with particular emphasis on supporting women in establishing careers in STEM fields, which addresses the clear bottleneck for women in research positions outlined in Fig. 1. For two specific examples of agriculture- and entomology-focused research institutes that have earned Athena SWAN accreditation, see the Roslin Institute (Scotland) and the Pirbright Institute (in England). Based on 10 key principles, the Charter was established in 2005 and originally focused on career progression for women in STEMM (STEM + medicine) fields; in 2015, the charter was expanded to address gender equality more broadly and recognized other disciplines (including the arts, humanities, social sciences, business and law), individuals in professional and support roles, and persons who are transgender. The Athena SWAN Charter has proven effective in a number of disciplines, and in order to apply for or receive competitive federal funding institutions must receive at least a bronze rank through the accreditation program. Linking accreditation programs like this to federal funding is a vital step in encouraging institutions to make hiring and retaining underrepresented scientists an integral part of their operations and ensures that resources are shifted to institutions that are achieving those goals. It forces institutions to move past traditional excuses and explore and implement lasting, perhaps creative, solutions.

In Australia, the Science in Australia Gender Equity (SAGE) pilot program is running the Athena SWAN accreditation through the Australian Academy of Science. The program launched its initial pilot period of two years in 2015, with 40 institutions including 30 Australian universities, six medical research institutes, and four government research institutions. This enrollment is notable considering there are 40 accredited public Australian universities; thus, the two-year pilot program includes 75% of Australian universities. Peer-review of the data collected during the pilot program will begin in mid-2017, and qualifying organizations will receive Athena SWAN awards in 2018. This is an impressive example, as the pilot program will produce considerable data in a short period of time.

Future Directions

Alignment with organizational diversity priorities, both locally, nationally, and internationally, should be a priority for institutions working to make STEM more inclusive. Funding bodies have made it clear they have expectations about the types of "broader impact" research should have on the community, and as researchers we should remember this is not government-mandated lip service. Visibility and representation are important, and when more young people and community members can see themselves in the scientists speaking to their communities the message will resonate with increased impact. Meaningful, community-led relationships are most productive, and building these networks is an activity that can incorporate trainees, junior staff, and the more senior members of a department.

Position descriptions should be more closely aligned with strategic plans at the departmental and organizational level. If increased engagement with the community is a goal, hires should be evaluated on their ability to conduct outstanding research and to create meaningful ways to build relationships with the public. Strategic plans communicate organizational goals for the future, and this is an effective way to create a bottom-up approach to ensuring milestones are met. Gender and other types of bias should be identified and addressed. Interestingly, men perceive research illuminating gender bias as having less merit than women do (Handley et al. 2015), so the most effective way to enact change remains to be determined.

Because of the many free or low-cost vehicles for digital knowledge transfer, expertise in online engagement should be something that organizational units begin to value highly. In the corporate sector, many people are employed with salaries that are not inconsequential to run public education and engagement campaigns. Increasing the "brand" recognition of a department, education campaign, or government department has a myriad of benefits, including that you build a network of like-minded folks who can help you advertise job openings, share relevant information, and raise the profile of your research across relevant geographic and disciplinary boundaries. When speaking with funding bodies and politicians, being able to quantitatively demonstrate the reach of your work is a real value add. Social media can also bring innovative collaborations together—for example, the authors of this manuscript met and learned of each other's work through Twitter. However, before creating yet another metric for researchers, it is important to recognize that in many departments this work is already being done, and what we truly lack is the adequate valuation of these activities in hiring and promotion.

Acknowledgments

MC Hardy is supported by a University of Queensland Research Fellowship (RM2013001889). We are grateful for insight on the topic provided on Twitter by Jenny Martin (@Jenny_STEM, Director of the Eskitis Drug Discovery Institute at Griffith University), Morgan Jackson (@BioInFocus, University of Guelph), Mark Carnall (@mark_carnall, Life Collections Manager at the Oxford University Museum of Natural History), and Matt Shipman (@ShipLives, North Carolina State University). The dataset generated and analyzed during the current study are available in Figshare: <https://doi.org/10.6084/m9.figshare.5027447.v1>.

References

Abramo, G., and C. A. D'Angelo. 2014. How do you define and measure research productivity? Scientometrics 101: 1129–1144.

Anonymous 2016. The trouble with GDP. The Economist.

Birdsall, C., S. Gershenson, and R. Zuniga. 2017. Stereotype threat, role models, and demographic mismatch in an elite professional school settings. IZA Discussion Paper No. 10459. January.

Boring, A., K. Ottoboni, and P. Stark. 2016. Student evaluations of teaching (mostly) do not measure teaching effectiveness. ScienceOpen Research.

Carpenter, T. A., N. Lagace, and S. Bahnmaier. 2016. Developing standards for emerging forms of assessment: the NISO altmetrics initiative. Ser. Librarian 70: 85–88.

Clancy, K.B.H., R. G. Nelson, J. N. Rutherford, K. Hinde, R. Lock, G. Raymond, P. Wheeler, L. Scott, M. Tobin, S. Marsham, et al. 2014. Survey of Academic Field Experiences (SAFE): trainees report harassment and assault. PLOS ONE 9: e102172.

Clauset, A., S. Arbesman, and D. B. Larremore. 2015. Systematic inequality and hierarchy in faculty hiring networks. Sci. Adv. 1: e1400005.

Constantine, M. G., L. Smith, R. M. Redington, and D. Owens. 2008. Racial microaggressions against Black counseling and counseling psychology faculty: a central challenge in the multicultural counseling movement. J. Couns. Dev. 86: 348–355.

Costanza, R., L. Fioramonti, and I. Kubiszewski. 2016. The UN sustainable development goals and the dynamics of well-being. Front. Ecol. Environ. 14: 59–59.

Currie, J., and B. Hill. 2013. Gendered universities and the wage gap: Case study of a pay equity audit in an Australian university. Higher Educ. Policy 26: 65–82.

Davidson, D. 1970. The furious passage of the black graduate student. Berkeley J. Sociol. 15: 192–211.

Ding, W. W., F. Murray, and T. Stuart. 2006. Gender differences in patenting in the academic life sciences. Science 313: 665–667.

Dunn, A. G., D. Surian, J. Leask, A. Dey, K. D. Mandl, and E. Coiera. 2017. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. Vaccine 35: 3033–3040.

Ewing, V. L., A. A. Stukas, and E. P. Sheehan. 2003. Student prejudice against gay male and lesbian lecturers. J. Social Psychol. 143: 569–579.

Eysenbach, G. 2011. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. J. Med. Internet Res. 13: e123.

Falk, M., and E. Long. 2016. Is physics open and accepting for LGBT people? APS News 25.

Fine, E., J. Sheridan, M. Carnes, J. Handelsman, C. Pribbenow, J. Savoy, and A. Wendt. 2014. Minimizing the influence of gender bias on the faculty search process, pp. 267–289. *In* V. Demos, C. W. Berheide, and M. T. Segal (eds.), Gender Transformation in the Academy, (Advances in Gender Research, Volume 19). Emerald Group Publishing Limited.

Funk, C., and B. Kennedy. 2016. The politics of climate change in the United States. Washington, DC.

Gentz, M. C., D. Sing, K. Gomes, T. DeFries, K. Manuel, and J. Grace. 2005. An educational program for Native Hawaiian high school students to stimulate interest in higher education in science and careers in entomology (poster). *In* Entomological Society of America Annual Meeting. 14–8 December, Fort Lauderdale, FL.

Ghiasi, G., V. Lariviere, and C. R. Sugimoto. 2015. On the compliance of women engineers with a gendered scientific system. PLoS ONE 10: e0145931.

Giannetti, B. F., F. Agostinho, C.M.V.B. Almeida, and D. Huisingh. 2015. A review of limitations of GDP and alternative indices to monitor human wellbeing and to manage eco-system functionality. J. Cleaner Prod. 87: 11–25.

Glass, J. L., S. Sassler, Y. Levitte, and K. M. Michelmore. 2013. What's so special about STEM? A comparison of women's retention in STEM and professional occupations. Social Forces 92: 723–756.

Grunspan, D. Z., S. L. Eddy, S. E. Brownell, B. L. Wiggins, A. J. Crowe, and S. M. Goodreau. 2016. Males under-estimate academic performance of their female peers in undergraduate biology classrooms. PLoS ONE 11: e0148405.

Handley, I. M., E. R. Brown, C. A. Moss-Racusin, and J. L. Smith. 2015. Quality of evidence revealing subtle gender biases in science is in the eye of the beholder. Proc. Natl. Acad. Sci. USA. 112: 13201–13206.

Hardy, M. C. 2014. Sexual harassment: create ethics codes to curb sex abuse. Nature 512: 136.

Hardy, M. C. 2016. Drafting an effective ethical code of conduct for professional societies: A practical guide. Admin. Sci. 6: 16.

Hardy, M. C., A. Carter, N. Bowden, A. G. Barnett, S. Bell, Á. Borrego, D. Drotar, D. Fanelli, F. M. Felisberti, R. Sear, et al. 2016. What do postdocs need to succeed? A survey of current standing and future directions for Australian researchers. Palgrave Commun. 2: 16093.

Haustein, S., T. D. Bowman, and R. Costas. 2016. Interpreting "altmetrics": viewing acts on social media through the lens of citation and social theories, pp. 372–406. *In* C. R. Sugimoto (ed.), Theories of informetrics and scholarly communication. De Gruyter, Berlin, Boston.

Irwin, J. 2002. Discrimination against gay men, lesbians, and transgender people working in education. J. Gay Lesbian Social Ser. 14: 65–77.

Krome, L. R. 2016. Attracting women to STEM programs: the influence of goal-orientations and the use of gendered wording in recruitment materials. K-State Electronic Theses, Dissertations, and Reports: 2004. Kansas State University, Kansas.

Lariviere, V., V. Kiermer, C. J. MacCallum, M. McNutt, M. Patterson, B. Pulverer, S. Swaminathan, S. Taylor, and S. Curry. 2016. A simple proposal for the publication of journal citation distributions. bioRxiv.

MacNell, L., A. Driscoll, and A. N. Hunt. 2015. What's in a name: Exposing gender bias in student ratings of teaching. Innov. Higher Edu. 40: 291–303.

Miller, D. I., and J. Wai. 2015. The bachelorâ€™s to Ph.D. STEM pipeline no longer leaks more women than men: A 30-year analysis. Front. Psychol. 6.

Misra, J., and J. Lundquist. 2015. Diversity and the Ivory Ceiling. Inside Higher Ed.

Misra, J., J. Lundquist, E. D. Holmes, and S. Agiomavritis. 2010. Associate Professors and Gendered Barriers to Advancement. Amherst, MA.

Misra, J., J. H. Lundquist, E. Holmes, and S. Agiomavritis. 2011. The ivory ceiling of service work. Academe. January.

Moss-Racusin, C. A., J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman. 2012. Science faculty's subtle gender biases favor male students. Proc. Natl. Acad. Sci. USA. 109: 16474–16479.

National Center for Education Statistics 2016. The Condition of Education - Postsecondary Education - Postsecondary Environments and Characteristics - Characteristics of Postsecondary Faculty - Indicator May (2016). Washington, DC.

Niemann, Y. F. 1999. The making of a token: a case study of stereotype threat, stigma, racism, and tokenism in academe. Front. J. Women Stud. 20: 111.

O'Kelly, F., G. J. Nason, R. P. Manecksha, S. Cascio, F. J. Quinn, M. Leonard, M. A. Koyle, W. Farhat, and M. J. Leveridge. 2017. The effect of social media (#SoMe) on journal impact factor and parental awareness in paediatric urology. J. Pediatric Urol. (in press).

Oakley, P., and C. Garforth. 1985. Guide to extension training. Food and Agriculture Organization of the United Nations.

OConnell, S., and M. A. Holmes. 2015. Hiring a diverse faculty, pp. 95–108. *In* M. A. Holmes, S. OConnell, and K. Dutt (eds.), Women in the Geosciences: Practical, positive practices toward parity. American Geophysical Union and John Wiley & Sons, Washington, DC.

Peoples, B. K., S. R. Midway, D. Sackett, A. Lynch, P. B. Cooney, M. Falagas, E. Pitsouni, G. Malietzis, G. Pappas, C. Alcaraz, et al. 2016. Twitter predicts citation rates of ecological research. PLoS ONE 11: e0166570.

Reuben, E., P. Sapienza, and L. Zingales. 2014. How stereotypes impair women's careers in science. Proc. Natl. Acad. Sci. USA. 111: 4403–4408.

Rodriguez, J., and K. J. Freeman. 2016. "Your focus on race is narrow and exclusive:" the derailment of anti-racist work through discourses of intersectionality and diversity. Whiteness Educ. 1: 69–82.

Russ, T., C. Simonds, and S. Hunt. 2002. Coming out in the classroom… an occupational hazard? The influence of sexual orientation on teacher credibility and perceived student learning. Commun. Educ. 51: 311–324.

Schnitzler, K., N. Davies, F. Ross, and R. Harris. 2016. Using Twitter to drive research impact: A discussion of strategies, opportunities and challenges. Int. J. Nursing Studies 59: 15–26.

Seglen, P. O. 1997. Why the impact factor of journals should not be used for evaluating research. BMJ 314: 498–502.

Sheltzer, J. M., and J. C. Smith. 2014. Elite male faculty in the life sciences employ fewer women. Proc. Natl. Acad. Sci. USA. 111: 10107–10112.

Smith, J. L., I. M. Handley, A. V. Zale, S. Rushing, and M. A. Potvin. 2015. Now Hiring! Empirically testing a three-step intervention to increase faculty gender diversity in STEM. BioScience 65: 1084–1087.

Stark, P., and R. Freishtat. 2014. An evaluation of course evaluations. ScienceOpen Research.

Symonds, M.R.E., N. J. Gemmell, T. L. Braisher, K. L. Gorringe, and M. A. Elgar. 2006. Gender differences in publication output: towards an unbiased metric of research performance. PLoS ONE 1: e127.

Turner, C.S.V., J. C. Gonzalez, and J. L. Wood. 2008. Faculty of color in academe: What 20 years of literature tells us. J. Divers. Higher Educ. 1: 139–168.

Upton, R., and C. Tanenbaum. 2014. The role of historically black colleges and universities as pathway providers: Institutional pathways to the STEM PhD Among Black Students. Washington, DC.

Wang, L. 2016. LGBT chemists seek a place at the bench. Chem. Eng. News 94: 18–20.

West, J. D., J. Jacquet, M. M. King, S. J. Correll, and C. T. Bergstrom. 2013. The role of gender in scholarly authorship. PLoS ONE 8: e66212.

Whittaker, J. A., and B. L. Montgomery. 2014. Cultivating institutional transformation and sustainable STEM diversity in higher education through integrative faculty development. Innov. Higher Educ. 39: 263–275.

Wille, L., and E. Derous. 2017. Getting the words right: when wording of job ads affects ethnic minorities' application decisions. Manage. Commun. Qtly. 89331891769988.

Yoder, J. D. 1991. Rethinking tokenism: Looking beyond numbers. Gender Soc. 5: 178–192.

Zambrana, R. E., R. Ray, M. M. Espino, C. Castro, B. Douthirt Cohen, and J. Eliason. 2015. "Don't Leave Us Behind:" the importance of mentoring for underrepresented minority faculty. Am. Educ. Res. J. 52: 40–72.