



# Hack weeks as a model for data science education and collaboration

Daniela Huppenkothen<sup>a,b,c,d,1</sup>, Anthony Arendt<sup>d,e</sup>, David W. Hogg<sup>c,b,f,g</sup>, Karthik Ram<sup>h,i</sup>, Jacob T. VanderPlas<sup>d</sup>, and Ariel Rokem<sup>d</sup>

<sup>a</sup>Institute for Data-Intensive Research in Astrophysics and Cosmology, Department of Astronomy, University of Washington, Seattle, WA 98195; <sup>b</sup>Center for Data Science, New York University, New York, NY 10003; <sup>c</sup>Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003; <sup>d</sup>The University of Washington eScience Institute, The Washington Research Foundation Data Science Studio, University of Washington, Seattle, WA 98105; <sup>e</sup>Polar Science Center/Applied Physics Laboratory, University of Washington, Seattle, WA 98105-6698; <sup>f</sup>Max-Planck-Institut für Astronomie, D-69117 Heidelberg, Germany; <sup>g</sup>Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010; <sup>h</sup>Berkeley Institute for Data Science, University of California, Berkeley CA 94720; and <sup>i</sup>Berkeley Initiative in Global Change Biology, University of California, Berkeley CA 94720

Edited by Russell A. Poldrack, Stanford University, Stanford, CA, and accepted by Editorial Board Member Marlene Behrmann July 9, 2018 (received for review September 29, 2017)

**Across many scientific disciplines, methods for recording, storing, and analyzing data are rapidly increasing in complexity. Skillfully using data science tools that manage this complexity requires training in new programming languages and frameworks as well as immersion in new modes of interaction that foster data sharing, collaborative software development, and exchange across disciplines. Learning these skills from traditional university curricula can be challenging because most courses are not designed to evolve on time scales that can keep pace with rapidly shifting data science methods. Here, we present the concept of a hack week as an effective model offering opportunities for networking and community building, education in state-of-the-art data science methods, and immersion in collaborative project work. We find that hack weeks are successful at cultivating collaboration and facilitating the exchange of knowledge. Participants self-report that these events help them in both their day-to-day research as well as their careers. Based on our results, we conclude that hack weeks present an effective, easy-to-implement, fairly low-cost tool to positively impact data analysis literacy in academic disciplines, foster collaboration, and cultivate best practices.**

data science | education | interdisciplinary collaboration | reproducibility

As data become cheaper to gather and store, researchers have become increasingly reliant on computational workflows requiring skills in statistical modeling, machine learning, and scalable computation. In addition, recent concerns about reproducibility crises motivate the acquisition of skills in open science and the design of reproducible workflows (e.g., refs. 1 and 2). Formal university curricula have been relatively slow to offer courses in these important topics, and this vacuum is often filled by extracurricular, ad hoc, less formal workshops. Well-known examples include software and data carpentry workshops, which provide training in research computing through a volunteer instructor program (3, 4). Meanwhile, there is an increase in statistical and computational courses designed for specific scientific disciplines, such as the Summer School in Statistics for Astronomers ([astrostatistics.psu.edu/su16/](http://astrostatistics.psu.edu/su16/)), the Google Earth Engine User Summits (<https://events.withgoogle.com/google-earth-engine-user-summit-2017/>), as well as a variety of project-focused (rather than pedagogical) meetings, such as the dotAstronomy meetings ([dotastronomy.com](http://dotastronomy.com)). Shorter meetings are also held in conjunction with conferences, such as the Hack Days at the annual American Astronomical Society meetings, the Brainhack hackathons associated with the meetings of the Organization for Human Brain Mapping and the Society for Neuroscience (5), and a hackathon at the American Geophysical Union meeting ([onlinelibrary.wiley.com/doi/10.1002/2014EO480004/pdf](http://onlinelibrary.wiley.com/doi/10.1002/2014EO480004/pdf)). In general, many of these events either tend to emphasize more traditional pedagogical class and lecture methodologies or focus on ad hoc projects developed during the

event (Fig. 1). Pedagogically focused events follow a classic academic model where novices learn new skills from experts. This model tends to focus on a one-way flow of information from instructor to student and is usually targeted toward participants in the training phase of their career. On the other end of the spectrum, project-focused workshops emphasize collaborative activities using existing skills, leading to the common perception that they are designed for technical experts. This may limit their audience. To bridge this gap, we describe here a model that we have implemented: “Hack Weeks” that aim to capitalize on the advantages of each of these models. These week-long events combine structured periods focused on pedagogy (often with an emphasis on statistical and computational techniques) and less structured periods devoted to hacks and creative projects, with the goal of encouraging collaboration and learning among people at various stages of their career.

We have run eight such hack week events: four focused on astronomy and two each focused on neuroscience and geoscience. Here we share the philosophy behind the hack week model, results from surveys of participants, practical lessons we have learned in organizing these events, and recommendations for future hack weeks. *SI Appendix* provides additional details on the practical aspects of organizing these events.

## Significance

**As scientific disciplines grapple with more datasets of rapidly increasing complexity and size, new approaches are urgently required to introduce new statistical and computational tools into research communities and improve the cross-disciplinary exchange of ideas. In this paper, we introduce a type of scientific workshop, called a hack week, which allows for fast dissemination of new methodologies into scientific communities and fosters exchange and collaboration within and between disciplines. We present implementations of this concept in astronomy, neuroscience, and geoscience and show that hack weeks produce positive learning outcomes, foster lasting collaborations, yield scientific results, and promote positive attitudes toward open science.**

Author contributions: D.W.H., K.R., J.T.V., and A.R. designed research; D.H., A.A., J.T.V., and A.R. performed research; D.H., A.A., J.T.V., and A.R. analyzed data; and D.H., A.A., D.W.H., K.R., J.T.V., and A.R. wrote the paper.

The authors declare no conflicts of interest.

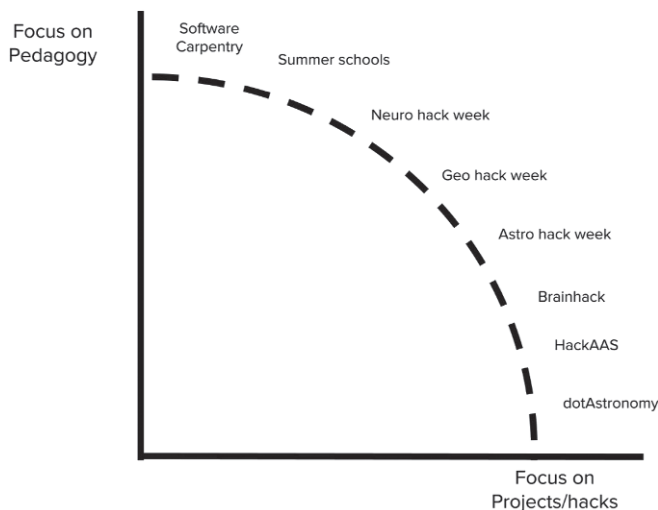
This article is a PNAS Direct Submission. R.A.P. is a guest editor invited by the Editorial Board.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

<sup>1</sup>To whom correspondence should be addressed. Email: dhuppenk@uw.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717196115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717196115/-DCSupplemental).

Published online August 20, 2018.



**Fig. 1.** Different types of events lie on a spectrum between an emphasis on pedagogy (e.g., Software Carpentry workshop) and an emphasis on project-based/hack-based activities (e.g., at science-oriented hackathons). Hack weeks also vary in the degree of emphasis on projects (e.g., Astro Hack Week, AHW) or pedagogy (e.g., Neuro Hack Week, NHW).

### What Is a Hack Week?

Our hack weeks combine structured, tutorial-style instruction with open-ended project work, providing opportunities for peer learning, networking, and building collaborations. In a space spanned by pedagogical focus as one dimension and focus on project work as the other, the hack weeks we have organized are designed to lie somewhere in between traditional summer schools and hackathons, where we believe they fill a space not currently fully addressed by existing models (Fig. 1).

The hackathon, a time-bounded, collaborative event that brings together participants around a shared challenge or learning objective (6), forms one primary axis of our events. Hackathons originated from the open-source software movement and have historically focused on software and technology development. In recent years, hackathons have evolved into a model providing opportunities for intensive, interdisciplinary collaboration (7) and education (8, 9) in the sciences. Core elements of hackathons include opportunities for networking, strengthening social ties, and building community connections, both within and across disciplines. Building on these core elements, hackathons have been implemented in different ways depending on the overall purpose, mode of participation, style of work environment, and participant motivation (10).

Summer schools have been designed to excel in transfer of knowledge from experts in the field to (early-career) researchers: They often serve as an entry point for scientists who aim to expand their research into a new area or switch fields. They are excellent at giving participants a reasonably deep understanding of a topic or field in a short amount of time. Within this concept, learning can take many forms, including traditional lecture formats but also hands-on project work, often in teams (e.g., Advanced Course on Computational Neuroscience, Okinawa Computational Neuroscience Course, Woods Hole Computational Neuroscience Summer course).

Our hack weeks extend the scientifically focused, communal hackathon model into a space that includes a strong element of pedagogy and peer learning. They aim to synthesize different goals and strategies from both models: They are more participant-driven than a summer school but have a stronger focus on pedagogy than a hackathon. Where a summer school is often organized around a framework of lectures and tutorials known in advance, hack weeks leave the majority of time to be

designed by participants, under careful facilitation of the organizers. Tutorials at hack weeks often serve as an entry point into a topic for further exploration and learning.

Hack weeks uniquely allow organizers to tailor the content of the workshop to the needs of the participants in an ad hoc fashion, including the number and content of tutorials. This way, the group as a whole can respond quickly or react to unforeseen challenges and opportunities. They encourage participants to self-organize in many different forms: experts working with other experts, mentoring relationships between experts and non-experts, or study groups among nonexperts, to name but a few. Hack weeks also allow participants to experiment with projects and ideas beyond their day-to-day research: For example, our hack weeks explicitly encourage projects around outreach and work aimed at improving the scientific community itself.

There is, however, a major risk in the lack of focus: By wanting to do many things at once, a hack week might potentially not do any of them well. Because tutorials are not necessarily the major focus of a hack week, the knowledge gained by participants in these tutorials may be shallow. A hack week carries a much larger risk of failure if objectives and expectations are not set by organizers well in advance, and clearly communicated to participants, because they often require significant preparation from the side of participants. Because of these risks, organizers face a much larger degree of uncertainty and need to be prepared to focus much of their energy on thoughtful selection and management of participants and facilitation of the large range of different types of activities at any given time (see also *SI Appendix, Section 4.2.4*).

We note that the terminology for these events is constantly evolving and that the “hackathon” concept may have implicit connotations that are disfavored in some communities. We also note that all of these events live on a constantly changing continuum, depending on the requirements of the scientific domain within which they live. For example, NHW is moving toward a more traditional summer school model, while AHW has strengthened its focus on projects and hacks in recent iterations.

### Why Run a Hack Week?

**Education and Training.** While some hack weeks are more focused on education than others (see Fig. 1), skill development in the form of tutorials as well as informal and peer learning is often a component. Furthermore, lateral knowledge transfer (3) through collaboration provides an opportunity to learn skills that are not described in papers and software implementations.

**Tool Development.** Hack weeks present an opportunity for scientific software developers to meaningfully engage with users and critically evaluate applications to particular scientific issues.

**Community Building.** Hack weeks are an opportunity to catalyze community development through a shared interest in solving computational challenges with open source software. They allow computationally minded researchers to break from the isolation of their institutions and spark new collaborations.

**Interdisciplinary Research.** Intensive, time-bounded collaborative events are an opportunity to experiment with concepts, questions, and methods that span boundaries within and across disciplines. Despite the fact that interdisciplinary experiments are impactful (11), they are often discouraged in traditional academic environments (12).

**Recruitment and Networking.** Hack weeks are a melting pot of participants from academia, government, and industry and provide numerous opportunities for networking. Close collaboration in diverse groups exposes skills that might be suitable for careers outside of a narrow domain.

**It Is Fun.** Hack weeks provide a respite from routine and a low-stress venue to learn new skills and attempt high-risk projects.

Note that the reasons for participants to attend a hack week are as diverse as the reasons for running such an event. Beginner participants may attend primarily to learn a new technique, while others may attend to gain experience in mentoring or to focus on an existing project already in progress (for more details on setting objectives, see *SI Appendix, Section 4.1.2*).

### Audience and Participant Selection

Hack weeks differ from many traditional conferences or summer schools in that knowledge transfer occurs across many levels of seniority and disciplinary boundaries. In addition, a substantial amount of hack week content is generated during the event itself, requiring active participation from attendees. In our experience, maximizing learning outcomes and collaborative exchanges at hack weeks requires a participant group that is diverse across categories of minority status, geographical origin, gender, discipline, and career stage, among others.

Traditional selection processes that rely heavily on internal heuristics of reviewers, especially those that consider characteristics peripheral to the evaluation criteria, are often fraught with personal and structural biases (e.g., ref. 13). To maximize diversity and minimize bias, we advocate for a selection process that is as quantitative and transparent as possible (13), enabling participants to hold organizers accountable for their selection decisions. This requires laying out a definition of successful participation, defining what criteria must be met to maximize the likelihood of success, and defining how those criteria will be assessed given the information about the candidates selected during the application stage.

For hack weeks, prerequisites will depend on the objectives of the workshop and may not exist at all. For example, AHW has traditionally accepted participants at all skill levels with respect to data science and did not include a merit-based selection, whereas NHW did include skill-based criteria in their selection (see also *SI Appendix, Section 4.1.5* for more detail on the individual selection procedures).

If merit-based selection is part of the evaluation process, organizers face the crucial decision of whether to assess merit blinded to other applicant characteristics. Because human decision makers tend to be swayed by unrelated characteristics including name (14) or gender (15), an initial merit selection blinded to demographic characteristics can be an effective way to counteract certain biases. A merit selection could then be performed via scores given independently by members of the organizing committee based on a set of predefined, explicit selection criteria. This type of blinded procedure tends to reduce biases when committees would otherwise not consider diversity during their selection (16).

However, a blind selection based purely on an assessment of merit will be counterproductive if it excludes participants who might have had less exposure to certain technologies or fewer opportunities to learn certain skills: For example, requiring a minimum level of programming experience will likely disadvantage candidates who have had fewer opportunities to learn programming due to structural inequalities. Additionally, blinding has been found to have negative effects on diversity for committees that already have a strong commitment to diversity, because these committees often correct for structural inequalities by considering demographic variables during merit selection (17). In this case, it may be beneficial to construct selection criteria that explicitly consider diversity and inclusivity (as NHW has done; see also *SI Appendix, Section 4.1.5*). Because systemic biases likely also enter at the application stage (where underrepresented groups may be less likely to apply), organizers should consider oversampling traditionally disenfranchised groups compared with the population of applicants.

No matter the selection procedures used, we encourage organizers to critically examine their cohort selection, experiment with new approaches, and routinely evaluate their procedure. For example, comparing demographic characteristics of the selected versus nonselected groups can unveil unintended biases during the merit-selection phase and thus allows adjustments in the procedure to mitigate or fully remove these effects.

### Themes

To date, all hack weeks we have organized have been subject-specific—that is, aimed at bringing together a community with a shared scientific interest, such as neuroscience. Advantages of this approach include shared language and scientific objectives within communities organized by subject, leaving more time for active collaboration on cutting-edge science. On the other hand, homogeneity may lead to group think and inhibit new, creative solutions. In this case, it may be advantageous to design a hack week around a technique (e.g., Gaussian Processes) or modality (e.g., imaging), such as the ImageXD (image processing across domains; [www.imagexd.org/](http://www.imagexd.org/)) meetings. For these events, building a shared vocabulary and shared understanding of major data analysis problems is crucial, but they also allow for cross-disciplinary diffusion of techniques into other subjects and therefore decrease the risk of duplication of method development efforts.

### Design Considerations

Several design elements contribute to the success of a hack week (see also *SI Appendix, Sections 4.1 and 4.2* for practical guidance). For example, scheduling: Longer events allow for a larger taught component, more ambitious projects, and cross-disciplinary exchanges. By spending more time together, participants are more likely to overcome barriers of professional terminology. But events that are too long may also lead to fatigue, resulting in a drop in positive outcomes later in the workshop. A well-designed hack week will have a clear schedule, limiting the number of parallel sessions and balancing the duration of taught components and project work.

The space used is also an important consideration. A hack week requires a flexible workspace that allows re-configuration, accommodating both lectures combined with interactive exchanges and individual work on laptops, as well as project work in small teams (see *SI Appendix, Sections 4.1.3 and 4.1.4*). Fortunately, many universities are experimenting with new types of spaces that allow these kinds of activities, and the adoption of active learning teaching methods (18) has led to the development of modular classrooms, designed for group activities and flexible seating arrangements.

Another design consideration is group size (see *SI Appendix, Section 4.1.4*). In a large group, chances for random participant exchanges may be reduced, and knowledge transfer may decline as the workshop fractures into smaller groups, often among participants who already know each other. If the group is small, participant selection must be especially carefully managed to achieve the desired level of diversity among participants to foster new collaborations. We have found groups with sizes between 40 and 70 participants to work well in enabling a breadth of projects while allowing the workshop to function as a cohesive group, but we encourage organizers to experiment with group size.

Hack week outcomes depend strongly on the interest and engagement of participants. Some attendees arrive with the goal of writing a scientific article; others plan to learn a specific topic (e.g., machine learning) or to analyze specific data using the tools covered in the tutorials. This leads to a wide variety of project types from sandbox-style explorations to focused work efforts. This breadth of possible outcomes makes it difficult to design for all possible participant goals and calls for adaptive, flexible design. The large variety in participant backgrounds and

experiences—and the resulting range in personalities and objectives of attendees—requires careful, active facilitation of both taught and project components of a hack week (see *SI Appendix, Section 4.2.4*). A well-advertised and enforced Code of Conduct is a very effective tool for managing expectations about participant interactions (see *SI Appendix, Section 4.1.7*). Community building is a core component of a hack week, and facilitation efforts need to consider both very strong personalities and very shy participants. In particular, the presence of impostor syndrome experienced by many participants must be taken into account during workshop design (see also *SI Appendix, Section 4.2.5* for concrete suggestions).

## Results

Measuring the success of a hack week objectively is complicated by the variety of goals that a hack week might have (see above). Additionally, the participant-driven format facilitates knowledge transfer and collaborations in sometimes surprising ways that escape traditional measures of success.

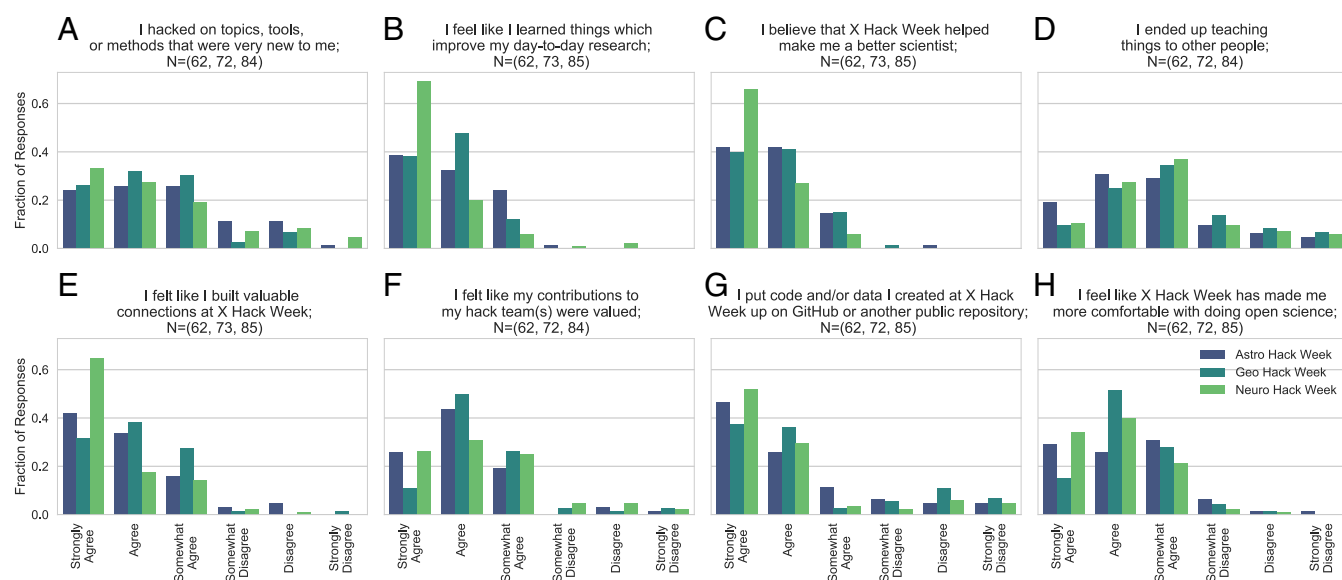
One key metric is the number of publications that result from hack week projects, but this is a fairly narrow definition of success, in line with standard academic performance indicators. Assuming that participants work largely in the open during a hack week and that most projects have a strong programming component, another indicator of success is the activity of participants in terms of code written and committed to a public code repository. Still, these measures ignore learning, community-building, as well as networking outcomes, which can be assessed through postworkshop surveys. Here, we have taken an approach that combines these metrics: We start with survey results and anecdotally report about publications and projects generated (see the following section).

Focusing on the outcomes of AHW, GHW, and NHW from 2016–2017, we find that most participants self-reported successful learning outcomes (AHW 76%, GHW 89%, and NHW 79% for responses “somewhat agree,” “agree,” and “strongly agree”; Fig. 2A). The overwhelming majority of respondents at the hack weeks (>95% for all events) believed that they learned things that improved their day-to-day research and that attendance has made them a better scientist (Fig. 2B and C). Because peer

learning is a major mode of knowledge transfer at hack weeks, we asked participants whether they taught other participants. We find that again a majority agreed with this statement to some degree (AHW 79%, GHW 69%, and NHW 75%; Fig. 2D), though responses are not as unequivocal as they are in some of the other categories. The majority of participants felt that they built valuable connections to other researchers (Fig. 2E), especially at NHW, where more than 64% of participants strongly agreed with this statement.

Given the diversity in skills and backgrounds of participants admitted to our hack weeks, a dependence of learning outcomes and teaching on career stage is plausible. At the same time, as suggested earlier, peer learning is a major mode of knowledge transfer for participants of all career stages. We find no strong evidence for a significant difference between early-career and senior participants in any of the hack weeks for questions regarding learning outcomes of new tools and topics (Fig. 2A), improvements in day-to-day research (Fig. 2B), or overall improvements in science (Fig. 2C) [ $p > 0.0007$  (trial-corrected significance threshold) with zero or small effect size for all]. Data and all (including nonsignificant) correlations are presented in *SI Appendix, Section 2, Table S1* and *SI Appendix, Section 3, Figs. S1–S12*. Only for GHW do we find that early-career researchers agree more strongly that the hack week improved their day-to-day research ( $P = 0.02$ ) with a large effect size of  $\phi_c = 0.35$  for 1.88 degrees of freedom (following ref. 19). Similarly, we find no indication ( $p > 0.0007$ ) that late-career researchers self-report a higher level of teaching at hack weeks compared with early-career researchers (Fig. 2D). However, the confidence intervals on the measured effect sizes are wide. Testing for the absence of an effect using an equivalence test on the effect size with an equivalence bound corresponding to a moderate effect size suggests that the data are currently not conclusive enough to reject a moderate to large effect ( $p_{eq} > 0.0007$  for all four questions above).

One important question is whether participants from underrepresented groups thrive at hack weeks or whether their full participation is impeded. Significant differences between minorities and nonminorities, even on a self-reported scale, for questions related to learning outcomes, teaching, or network



**Fig. 2.** Postworkshop survey responses from the 2016 AHW, GHW, and NHW. Response rates are in the panel titles. Results are presented in three different domains: the development of technical skills (A–C), collaboration and teaching (D–F), and shifts in attitudes toward reproducibility and open science (G and H).

building, would indicate that improvements in workshop facilitation and structure may be required to allow members of these groups to participate fully. In our surveys, we find no significant dependence of self-reported learning outcomes on gender identity or race/ethnicity ( $p > 0.0007$ ). Similarly, for none of the hack weeks do minority participants differ significantly in their answers with respect to teaching outcomes, building valuable connections, or the value of their contributions to their hack teams ( $p > 0.0007$ ). For GHW, there is an indication that participants from racial/ethnic minorities may respond more positively when asked about building connections ( $p = 0.04$ ) with a medium to large effect size ( $\phi_c = 0.4$ ;  $\text{dof}_{\phi_c} = 0.97$ ), while for AHW, responses regarding the value of contributions to hack teams may be spread more widely for participants from racial/ethnic minorities than for Caucasian participants ( $p = 0.01$ ;  $\phi_c = 0.4$ ;  $\text{dof}_{\phi_c} = 0.98$ ). Equivalence tests reveal systematically small, though nonsignificant,  $p$  values in the range of  $p = 0.02 - 0.05$  for AHW for all four questions above in conjunction with gender or ethnic/racial identity, while for GHW and NHW  $p > 0.05$  for the same questions (note that the sample size for GHW and NHW was smaller by a factor of 2 compared with AHW). These results, while not a decisive exclusion of an effect of race/ethnicity or gender on hack week participation, provide an indication that our facilitation strategies may be effective in fostering participation. Future work on how demographics interact with hack week attendance may be fruitful.

We find that the hack weeks have been largely successful in promoting positive attitudes toward reproducibility and open science: At all three events, the majority reported that the hack week has made them more comfortable with open science (GHW 97%, NHW 95%, AHW 72%; Fig. 2H), and more than 85% of all participants (AHW 86%, GHW 94%, NHW 95%; Fig. 2G) put code or data created at the hack week into a public repository. While the focus on open science is not necessarily a required component of a hack week, it aligns naturally with many of the goals and values commonly promoted at hack weeks, such as production of open-source software and data sharing. In some fields, especially where ethical issues around data sharing and privacy are relevant, this should be augmented with a discussion of ethical considerations.

In line with the surveys' exploratory nature, these results should be read only as an initial indication of the hypotheses we proposed about the use and outcomes of hack weeks. The number of respondents is small and the effects likely subtle, and lack of significant differences may be due to statistical power in our sample. Furthermore, the most important independent variable—attendance of a hack week—is not accounted for in our current design. Moreover, self-reported learning outcomes are not an objective measure, because they are likely subject to response biases. Future work will include more refined survey design and inclusion of control nonattendees.

Because all three events are relatively recent, it is still early to evaluate long-term outcomes, including publications and collaborations resulting from these events. There are, however, initial indicators that all hack weeks encouraged long-term engagement with new concepts or tools and that they directly resulted in a number of publications (20–27). Specific examples follow below.

### Examples of Hack Week Outcomes

**Example 1: AHW.** In 2015, a small team used AHW to found a new software project called Stingray (<https://github.com/StingraySoftware/stingray>) with the goal of providing implementations of time series analysis algorithms often used in astronomy. AHW enabled participants to seed a new collaboration around a software project needed by the larger community, facilitated by the collaborative environment at AHW. Stingray has since matured into an enduring collaboration within the com-

munity with five active maintainers and four Google Summer of Code projects.

**Example 2: GHW.** In 2016, a GHW project team used Google Earth Engine to explore spatial patterns in climate, topography, and population data with the goal of mapping the most suitable locations for renewable energy sites in the United States. The team used machine learning algorithms in conjunction with the powerful hardware resources provided by Google Earth Engine ([georgerichardson.net/2017/04/10/searching-for-energy-in-a-random-forest/](http://georgerichardson.net/2017/04/10/searching-for-energy-in-a-random-forest/)).

**Example 3: NHW.** During NHW 2016, one of the teams analyzed an openly available dataset of MRI data from children (28), to test the effects of motion on analysis results, using varying motion cutoffs. The team (all from different institutions) continued to work on this project remotely after the end of NHW, eventually publishing a paper describing these results (23).

### Conclusions

The fast-paced changes in the computational and methodological landscape require that traditional fields of science rapidly adapt to new data analysis challenges. To address these challenges, new types of workshops, including unconferences, hackathons, and bootcamps, have been developed in recent years in various scientific disciplines and now exist alongside with and support the existing structure of academic conferences, formal classes, and other learning opportunities. Here, we introduce one such concept, hack weeks, and detail the underlying philosophical ideas along with experiences from events held in three different fields.

Hack weeks serve multiple purposes, including dissemination of technological advances through the scientific community, building collaborations between academic subdisciplines, and fostering interdisciplinary research. Initial results from six events held in 2016 and 2017 in three different fields (astronomy, geosciences, and neurosciences) indicate that hack weeks succeed at all of these objectives.

Hack weeks are still a very young concept, and estimating the long-term impact of these events within the scientific communities they serve will require follow-up over multiple years, to assess their effect on collaboration networks, career outcomes, and adoption of new methods. We have shown, however, that hack weeks provide an easy-to-implement, fairly low-cost method to introduce new technologies and methods into scientific fields on much shorter time scales than traditional teaching efforts. While we focus here on hack weeks in scientific fields, the concept could be extended to other areas and is more generally useful in any area (*i*) where useful tools can be learned in short tutorials, (*ii*) where results and outcomes can be produced on the time scale of a few days, and (*iii*) that would benefit from collaborative approaches that cross traditional boundaries. Such areas could include the social sciences, the humanities, as well as music and art.

### Materials and Methods

We performed postattendance surveys for AHW, GHW, and NHW in 2016 and 2017. All surveys contained general questions about attitudes toward the workshop as well as open science and reproducibility, shared among all three surveys. Response rates for NHW (2016: 41 responses; 2017: 45 responses) and GHW (2016: 42 responses; 2017: 41 responses) were 100% in both years; the response rate for AHW was 71% (35 out of 49) in 2016 and 82% (37 out of 45) in 2017. Participants were asked to respond to statements regarding these topics using a six-point Likert-type scale. All questions were anonymously recorded. The experimental procedures were approved by the Institutional Review Boards at University of Washington; New York University; and University of California, Berkeley. All participants gave their informed consent. No responses were discarded, and no preprocessing was performed on the data. We test for correlations between

demographic characteristics (independent variable) and question responses (dependent variable) using a standard  $\chi^2$  test and compute the effect sizes via a bias-corrected version of Cramér's  $V$  (29, 30), denoted  $\phi_c$ . We additionally perform equivalence tests on the effect size to quantify the absence of correlations. The full procedure is available in *SI Appendix, Section 1* and online (see the repository: <https://github.com/uwescience/HackWeek-Writeup>).

**ACKNOWLEDGMENTS.** The authors thank the participants in the hack weeks that we have organized for their myriad contributions to this work. The authors thank the anonymous reviewers and the editor for their helpful comments and suggestions; Laura Norén for help on ethics and Institutional Review Board; Stuart Geiger for helping to formulate the survey;

Christine Huebner for advice on statistics; Brittany Fiore-Gartland, Laura Norén, and Jason Yeatman for comments on the manuscript; and Tal Yarkoni for advice regarding automated selection procedures. This work was partially supported by the Moore-Sloan Data Science Environments at University of California, Berkeley; New York University; the University of Washington; and the Washington Research Foundation Fund for Innovation in Data-Intensive Discovery. NHW is supported through National Institute of Mental Health Grant 1R25MH112480. D.H. is partially supported by the James Arthur Postdoctoral Fellowship at New York University and acknowledges support from the DIRAC Institute in the Department of Astronomy at the University of Washington. The Institute for Data-Intensive Research in Astrophysics and Cosmology is supported through generous gifts from the Charles and Lisa Simonyi Fund for Arts and Sciences and the Washington Research Foundation.

- Pashler H, Wagenmakers E-J (2012) Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspect Psychol Sci* 7:528–530.
- Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533:452–454.
- Wilson G (2014) Software carpentry: Lessons learned. *F1000 Res* 3:62.
- Teal TK, et al. (2015) Data carpentry: Workshops to increase data literacy for researchers. *Int J Digital Curation* 10:135–143.
- Craddock CR, et al. (2016) Brainhack: A collaborative workshop for the open neuroscience community. *Gigascience* 5:16.
- Decker A, Eiselt K, Voll K (2015) Understanding and improving the culture of hackathons: Think global hack local. *2015 IEEE Frontiers in Education Conference (FIE)* (IEEE Computer Society, Washington, DC), pp 1–8.
- Groen D, Calderhead B (2015) Science hackathons for developing interdisciplinary research and collaborations. *eLife* 4:e09944.
- Kienzler H (2015) Bringing students into research by hacking global health. *High Educ Res Netw J* 10:17–29.
- Lamers MH, Putten P, Verbeek FJ (2014) Observations on tinkering in scientific education. *Entertaining the Whole World*, Human-Computer Interaction Series (Springer, London), pp 137–145.
- Drouhard M, Tanweer A, Fiore-Gartland B (2017) A typology of hackathon events. *Hacking at Time-Bound Events Workshop at Computer Supported Cooperative Work 2016 (CSCW'16)* (ACM, New York), pp 1–4.
- Hall KL, et al. (2012) Assessing the value of team science: A study comparing center- and investigator-initiated grants. *Am J Prev Med* 42:157–163.
- Sung NS, et al. (2003) Science education. Educating future scientists. *Science* 301:1485.
- Sunstein CR, Hastie R (2015) *Wiser: Getting Beyond Groupthink to Make Groups Smarter* (Harvard Business Press, Brighton, MA).
- Bertrand M, Mullinathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Am Econ Rev* 94:991–1013.
- Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J (2012) Science faculty's subtle gender biases favor male students. *Proc Natl Acad Sci USA* 109:16474–16479.
- Bohnet I (2016) *What Works: Gender Equality by Design* (Harvard Univ Press, Cambridge, MA).
- Behaghel L, Crépon B, Barbanchon TL (2015) Unintended effects of anonymous resumes. *Am Econ J Appl Econ* 7:1–27.
- Prince M (2004) Does active learning work? a review of the research. *J Eng Educ* 93:223–231.
- Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Routledge, New York), 2nd Ed.
- Gully-Santiago M, Jaffe DT, White V (2015) Optical characterization of gaps in directly bonded Si compound optics using infrared spectroscopy. *Appl Opt* 54:10177–10188.
- Faria JP, et al. (2016) Uncovering the planets and stellar activity of CoRoT-7 using only radial velocities. *Astron Astrophys* 588:A31.
- Keshavan A, et al. (2017) Mindcontrol: A web application for brain segmentation quality control. *Neuroimage* 15:365–372.
- Leonard J, Flournoy J, de los Angeles CPL, Whitaker K (2017) How much motion is too much motion? Determining motion thresholds by sample size for reproducibility in developmental resting-state MRI. *Res Ideas Outcomes* 3:e12569.
- Jordan K, Keshavan A, Mandelli ML, Henry R (2017) Cluster-viz: A tractography QC tool. *Res Ideas Outcomes* 3:e12394.
- Peterson D (2017) Streamlining the process of 3d printing a brain from a structural MRI. *Res Ideas Outcomes* 3:e13394.
- Hahn C, et al. (2017) Approximate Bayesian computation in large-scale structure: Constraining the galaxy-halo connection. *Mon Not R Astron Soc* 469:2791–2805.
- Price-Whelan AM, Hogg DW, Foreman-Mackey D, Rix H-W (2017) The Joker: A custom Monte Carlo sampler for binary-star and exoplanet radial velocity data. *Astrophys J* 837:20.
- Di Martino A, et al. (2014) The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry* 19:659–667.
- Cramér H (1946) *Mathematical methods of statistics (Almqvist & Wiksells Akademiska Handböcker)* (Princeton Univ Press, Princeton, NJ).
- Bergsma W (2013) A bias-correction for Cramer's and Tschuprow's. *J Korean Stat Soc* 42:323–328.