Hoping for optimality or designing for inclusion: Persistence, learning, and the social network of citizen science

Julia K. Parrish[a,1], Timothy Jones[a], Hillary K. Burgess[a], Yurong He[a], Lucy Fortson[b], and Darlene Cavalier[c]

[a]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020; [b]School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455; and [c]School for the Future of Innovation in Society, Arizona State University, Tempe, AZ 85287-5603

The explosive growth in citizen science combined with a recalcitrance on the part of mainstream science to fully embrace this data collection technique demands a rigorous examination of the factors influencing data quality and project efficacy. Patterns of contributor effort and task performance have been well reviewed in online projects; however, studies of hands-on citizen science are lacking. We used a single hands-on, out-of-doors project—the Coastal Observation and Seabird Survey Team (COASST)—to quantitatively explore the relationships among participant effort, task performance, and social connectedness as a function of the demographic characteristics and interests of participants, placing these results in the context of a meta-analysis of 54 citizen science projects. Although online projects were typified by high (>90%) rates of one-off participation and low retention (<10%) past 1 y, regular COASST participants were highly likely to continue past their first survey (86%), with 54% active 1 y later. Project-wide, task performance was high (88% correct species identifications over the 31,450 carcasses and 163 species found). However, there were distinct demographic differences. Age, birding expertise, and previous citizen science experience had the greatest impact on participant persistence and performance, albeit occasionally in opposite directions. Gender and sociality were relatively inconsequential, although highly gregarious social types, i.e., "nexus people," were extremely influential at recruiting others. Our findings suggest that hands-on citizen science can produce high-quality data especially if participants persist, and that understanding the demographic data of participation could be used to maximize data quality and breadth of participation across the larger societal landscape.

citizen science | crowdsourcing | data quality | retention | dabblers

Citizen science, defined broadly as the free choice participation of the nonexpert public in the practice of science with the programmatic goal of scientific outcomes (1), is an rapidly expanding area of science (2, 3), science literacy (4), and science communication (5). Since its creation in 2009, the citizen science clearinghouse SciStarter has grown to more than 2,100 projects listed on the site (*SI Appendix*, Fig.S1), with the greatest absolute number and relative growth in hands-on projects, or those with activities requiring the physical presence and work of the project participants. Online citizen science, or projects conducted via the World Wide Web, is also growing. The number of projects hosted by the online citizen science platform Zooniverse has grown at a steadily increasing rate (*SI Appendix*, Fig. S1) to more than 140 projects in disciplines as diverse as astronomy and environmental conservation. This platform now accounts for the majority of online citizen science registrants in the world (1.75 million) (6). Collectively, hands-on and online citizen science translates into dozens of scientific advancements (measured as peer-reviewed publications) annually (7), and billions of in-kind dollars (8, 9). In the ideal, citizen science provides a way to achieve massive data collection and data processing needs while elevating the awareness, knowledge, and understanding of the nonscience public about the practice of science and the relevance of scientific outcomes.

Participant Persistence and Project Retention. Within citizen science, scientific and societal goals are dependent, at least in part, on the amount of time and effort that individual participants put toward a given project. Sauermann and Franzoni (8) point out that, even though the collective contribution of online volunteerism in crowdsourced image classification projects is vast, individual-level contributions follow a classic Pareto curve: most contributors complete only a single task, and relatively few (∼10%) contribute the majority (∼80%) of the work. Online projects are not the only type of citizen science displaying a steep fall-off in effort and retention. Boakes et al. (10) characterized participant contribution in sensor/sampling projects (defined here as hands-on projects requiring the participant to collect samples, often with specific sensor technology) as falling into three statistically derived volunteer engagement profiles: dabblers, or those who sample science lightly; enthusiasts, or those who perform the vast majority of the work; and steady volunteers, or those individuals in between. Dabblers formed the vast majority of the participant corps (67–84%) whereas enthusiasts were vanishingly rare (1–4%), even though this latter group accomplished a proportionally large fraction of the observations. Dabbling may thus define a large portion of the citizen science corps, although a comprehensive examination of hands-on projects is lacking.

Participant Performance and Data Quality. Whether nonexpert participants are successful in contributing to the science outcomes of a citizen science project is certainly dependent on the number of participants and how long they persist, but also on how accurate these contributors are at whatever task(s) they perform. One of the strongest criticisms of citizen science leveled by the mainstream science community is of a lack of rigor in data collection (11, 12). The relationship between how good a participant is at mastering a task and how long they have been

performing the task has been shown to be a function of the complexity of the task (8), the type of training (13), and the amount and type of feedback participants receive when they begin to collect data (14).

In online projects, individual performance or accuracy can be relatively inconsequential, as crowdsourcing (i.e., multiple, independent contributors or players performing the exact same task) across a large participant base can optimize results. For example, most online classification or transcription projects (e.g., projects in the Zooniverse) feature relatively simple tasks such as counting the number of, and/or identifying the type(s) of, objects in an image. These are easily mastered with short tutorials and minimal practice (15). When a task has been completed by the requisite number of contributors (usually 5–10, depending on project), algorithm voting routines can allow the return of a collective answer rivaling that of expert opinion (16). By contrast, in online gaming projects (e.g., Foldit), the majority of the participants are relegated to a nonessential role as expert players with "winning" or optimal solutions to the task at hand emerge from the crowd (17). In both of these designs, data returned by one-off dabblers can be useful in generating scientific outcomes (15).

By contrast, in hands-on citizen science, mastering the necessary skills to return high-quality data may take time, requiring repetitive task performance over multiple data collection sessions to achieve the necessary precision and/or accuracy (18). For some hands-on projects, the data collection accuracy of incoming participants may already be high. Hobbyists or amateur experts who have learned and honed their craft are attracted to and/or preferentially recruited into projects affording them demonstration of their prowess. Birders and amateur astronomers are exemplars (19). However, projects hoping to recruit a broader participant base must rely on participants persisting through to some threshold of performance after which the project can credibly use their data. Thus, there is a theoretical trade-off between the amount of time or effort required "in practice" to attain the minimum scientific standard of the project and the retention curve described by the participating population. If the majority of the recruited population are dabblers, these projects will not succeed in delivering the "science at scale" promised by citizen science.

Averages Do Not Describe the Everyman. From the perspective of science, citizen science should optimize retention and data collection accuracy. What attracts and retains participants is a central issue within citizen science (20, 21), as is whether and how participants learn (4, 19). At a fundamental level, alignment between the interests of the participant and the mission of the project undergirds the motivation to persist (22). Aspects of personal experience can predispose individuals toward participation, and may also correlate with subsequent performance, including but not limited to gender (23), age (24), hobbies (19), and the tendency to work with others or gregariousness (25). Whether any/all of these participant attributes influence performance, and specifically the relationship between persistence and learning, is unknown.

In this paper, we restrict the definition of citizen science to those projects explicitly delivering data and requiring the active mental participation of contributors (i.e., active citizen science) (26). We quantitatively explored the intersections among participant persistence, participant performance, and project design in active citizen science, with an in-depth analysis of a single hands-on, out-of-doors citizen science project, the Coastal Observation and Seabird Survey Team (COASST). We frame this case study with a meta-analysis of the retention characteristics of 54 citizen science projects across the hands-on, online continuum. Our analyses provide a comparison of contributor effort across the landscape of citizen science. Within hands-on projects, our work suggests that attempting to optimize retention, data

collection accuracy, and social connectedness selects against top performers in favor of a broader participant base. We discuss these findings in light of claims that citizen science can return high-quality science and participant learning and self-fulfillment for the majority of the nonexpert public (1, 3), thus providing a path toward the democratization of science (27).

Results

COASST is an active citizen science project recruiting residents of coastal communities along the northwest coast of the continental United States as well as Alaska (28). After an in-person, expert-led training session, attendees who sign up to participate begin to conduct standardized, effort-controlled surveys of "their" beach at least monthly, searching for beach-cast carcasses of marine birds. For all carcasses found, participants record meristic (e.g., number and arrangement of toes) and morphometric (e.g., wing measurement) data before using these same pieces of evidence to make a taxonomic identification with the aid of a project-specific field guide. All carcasses are uniquely tagged, photographed, and left in place. Taxonomic identifications are independently verified by experts via submitted measurement and photographic data.

The Retention Landscape. On average, 71% of COASST participants ($n = 2,511$) persisted after their first survey, and 40% continued past 1 y (Fig. 1A). If guests are excluded from this sample (i.e., only regular participants remain), retention past 1 y exceeded 50% (Fig. 1A), with 14% still active in the project 5 y later. Not surprisingly, these long-term participants sustained a larger fraction of total effort, measured as surveys performed ($R^2 = 0.39$; *SI Appendix*, Fig. S2A). However, this time/effort relationship broke down when effort was assessed as carcasses found ($R^2 = 0.09$), with a wide range of cumulative carcass counts across the participant population (*SI Appendix*, Fig. S2B). That is, many individuals persisted for years in surveying even when they repeatedly found few to no birds. Fully one third (34%) of long-term participants never found a carcass, and 11% found only one (*SI Appendix*, Fig. S2C). Thus, retention in COASST does not appear to be related to effort-based rewards (e.g., finding birds) more typical of hobbyist projects.

To put the COASST retention statistics into perspective, we explored the larger patterns of contributor effort defined by active citizen science projects reporting the percentage of participants who dabbled (defined as completing a single task, survey, or observation) and/or the percentage who were still active in the project 1 y later (Fig. 1B). We used literature- and project-reported values (project names and data sources are provided in *SI Appendix*, Table S1) to examine four project classes: (*i*) online crowdsourcing projects recruiting multiple participants to independently classify or transcribe digital information, (*ii*) online crowdsourcing projects inviting participants to compete for optimal solutions to a problem set within a gaming framework, (*iii*) hands-on projects asking participants to collect digital or real-world samples, and (*iv*) hands-on projects asking participants to collect deductive data, or conclusions based on in situ evidence, including COASST.

Online projects, and especially image classification projects, attracted a high proportion of dabblers, ranging from 50% to ≥90% of registered participants, and had uniformly low retention past 1 y (<10%). For some of these projects, participant turnover at 1 y was total. By contrast, hands-on projects had extremely wide ranges of dabbling and retention past 1 y (hereafter simply "retention"). Within this group, three subsets emerged, including high-dabbler, low-retention projects intensively recruiting for a single, local event (e.g., bioblitzes) (29). At the other extreme were low-dabbler, high-retention projects. These included hobbyist projects passively recruiting skilled amateurs (e.g., birders to Breeding Bird Surveys) (19, 30) as well as projects actively recruiting individuals with a high interest in the mission of the project (e.g., sense of place associated with local environmental monitoring) (31). Between these poles
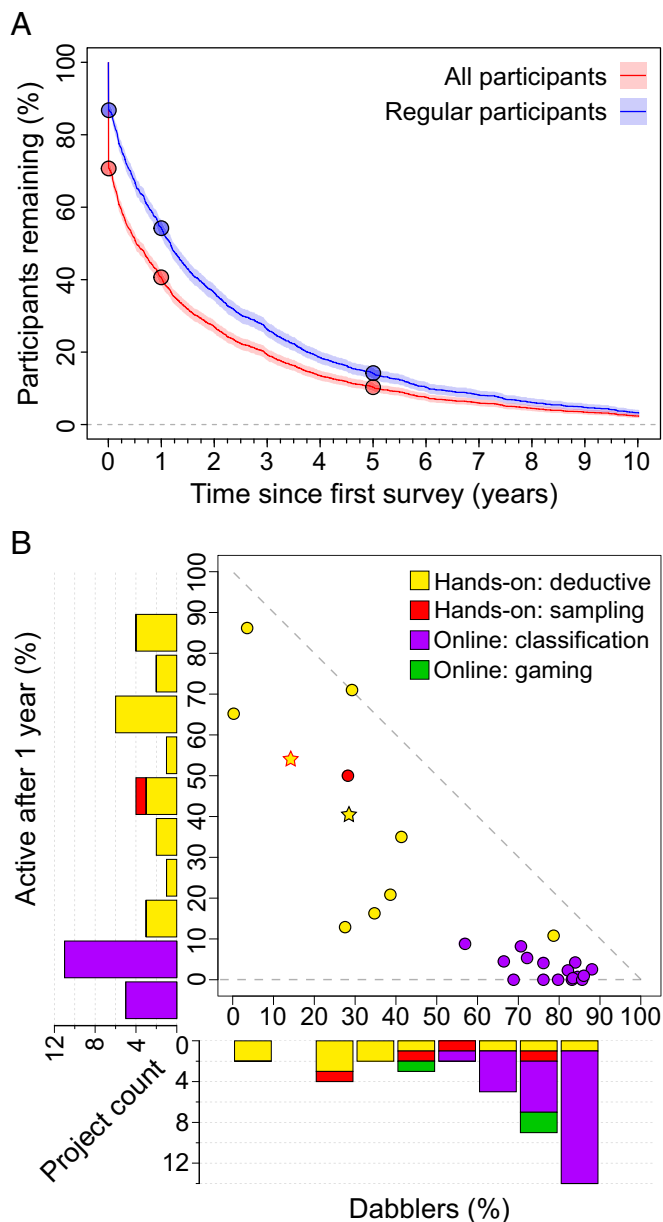
Fig. 1. The retention landscape in active citizen science. (*A*) Retention in the COASST program measured as the duration between first and last survey date for all participants (red; *n* = 2,511) and for regular participants only (i.e., subtracting guests; blue; *n* = 1,654). Points along the curve (left to right) indicate time points following the first survey, at 1 y, and at 5 y. (*B*) Results of a meta-analysis of 54 citizen science projects reporting one or both of the percentage of the contributing population interacting for only a single session, observation, task, or survey (i.e., dabblers; *n* = 41) and the percentage of the contributing population remaining active in the project past 1 y (*n* = 39). Projects with total turnover at 1 y are in the bar below zero. Histograms are binned counts of all projects for which the relevant data were obtainable; scatterplot points represent projects for which both metrics were available (*n* = 26). Projects were a priori categorized into four exclusive classes. Stars represent the COASST project (black outline, all participants; red outline, regular participants only). *SI Appendix*, Table S1 provides a list of projects and sources.

were a range of projects offering participants the chance to develop and hone activity-specific skills over multiple data collection opportunities, from national- to global-scale projects with Web-based training and feedback [e.g., Nature's Notebook (NN), Community Collaborative Rain Hail and Snow Network (CoCoRaHS), Reef

Environmental Education Foundation (REEF); *SI Appendix*, Table S1] to local- to regional-scale projects with in-person training and personalized feedback (e.g., Virginia Master Naturalists, COASST; *SI Appendix*, Table S1). Because all projects represented in the retention landscape have been successful over multiple years (even decades), and range from hundreds to hundreds of thousands of participants annually, our analysis suggests that fundamental differences may arise as a function of design, and secondarily implicate particular types of participants, at least in the highest-retention projects.

Practice Makes (Nearly) Perfect. Although COASST does not post the highest rates of retention (Fig. 1*B*), it also recruits a large proportion of self-defined beginners (31), making it a useful exemplar of the design tension between retention and data collection accuracy when the participant base is broad in terms of skill. The ability of COASST participants to correctly identify a carcass to species (hereafter "accuracy") is central to the experience of the individual and the data quality of the project (28). Becoming good requires practice. For participants who did encounter carcasses, it is clear that incoming participants (<10 carcasses found) displayed an extreme range of variability in identification skill (calculated as accuracy across all birds found by a given individual; Fig. 2*A*). However, participants logging higher numbers of carcasses quickly boosted their cumulative accuracy, and some individuals were truly exceptional (>95% correct).

Learning, operationally defined as an increasing likelihood of deducing the correct species of carcass encountered, happened continuously, from survey to survey and even from bird to bird. At the survey level, accuracy was affected by the prior experience of the survey team [model Akaike information criterion (AIC) = 8,973, D^2 = 10.1%; *SI Appendix*, Table S2], as well as by carcass abundance and taxonomic diversity, all of which displayed asymptotic relationships (*SI Appendix*, Fig. S3). In other words, the opportunity to sample more carcasses on a survey improved team performance, regardless of the inherent experience of its members (ΔAIC = 553; *SI Appendix*, Fig. S3*B* and Table S2). However, this effect was dampened by species diversity (ΔAIC = 367; *SI Appendix*, Fig. S3*C* and Table S2); that is, it was easier to identify 10 examples of the same species than single examples of 10 different species. Because many participants do the survey with partners (~50% of all surveys), we also modeled accuracy at the carcass level, at which the identification was credited only to the most experienced survey team member (Fig. 2*B*). The resultant "learning curve" had an intercept at 78% (95% CI, 75.4–80.5%) and an asymptote at 89% (95% CI, 88.2–89.6%), although achieving this level took time. For example, the average participant surveying Pacific Northwest outer coast beaches would be within 2% of the asymptote after ~43 carcasses or ~1.25 y (15 surveys).

Taken together, these analyses indicate that participants in COASST can become highly adept at species identification if they persist long enough. We modeled the shape of that trade-off "frontier" between persistence and accuracy by using Pacific Northwest outer coast participants as an example (Fig. 2*C*). If COASST required extremely high levels of accuracy (e.g., >90%) before participant data would be accepted, the project would collapse, as achieving this theoretical level of performance is beyond the duration of even the most persistent participant. Relaxing accuracy requirements to 75% would allow COASST to capture data from the entire participant population. Thus, knowledge of the shape of the trade-off frontier gives project managers information about what they are gaining (or losing) by applying a particular stricture and/or information on where to apply additional training to boost participant acumen.

All Participants Are Not the Same. Because individual participants are not described by project-level averages, we explored the
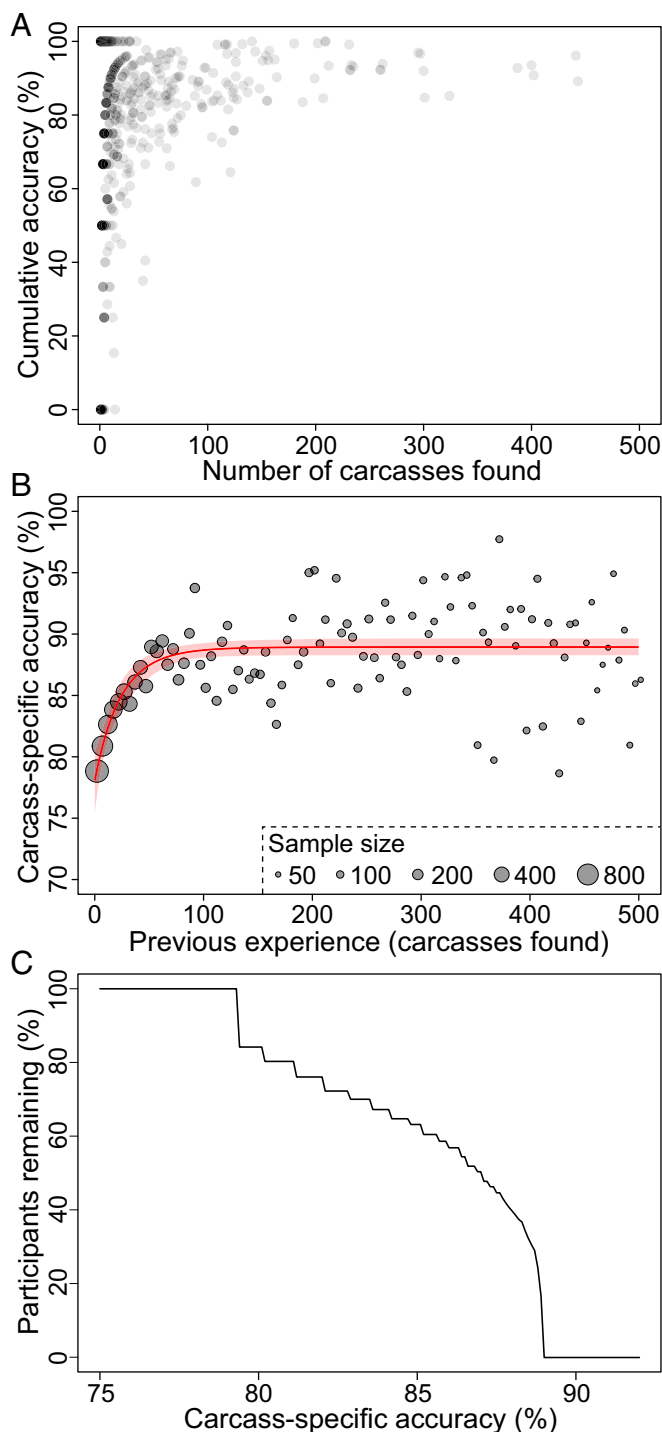
Fig. 2. (*A*) Scatterplot of cumulative participant data collection accuracy (i.e., percentage of carcasses identified correctly) of all individuals (*n* = 774; dots) as a function of the total number of carcasses they found. (*B*) The population-level learning curve (mean and 95% CI) for taxonomic identification (i.e., the likelihood of correctly identifying a particular carcass, or carcass-specific accuracy) as a function of participant experience (proxied by previous carcasses found by the most experienced team member) modeled as a negative exponential (based on 774 participants, as in *A*, and 19,831 birds). Dots represent raw data binned (bin size = 5) as a function of experience, with dot size corresponding to the number of carcasses in each bin (i.e., more carcasses were found by less experienced people, on average, than by participants who had found at least 400 carcasses). (*C*) Trade-off frontier between participant retention and accuracy tuned to Pacific Northwest outer coast sites (average annual carcass encounter rate of 30 birds).
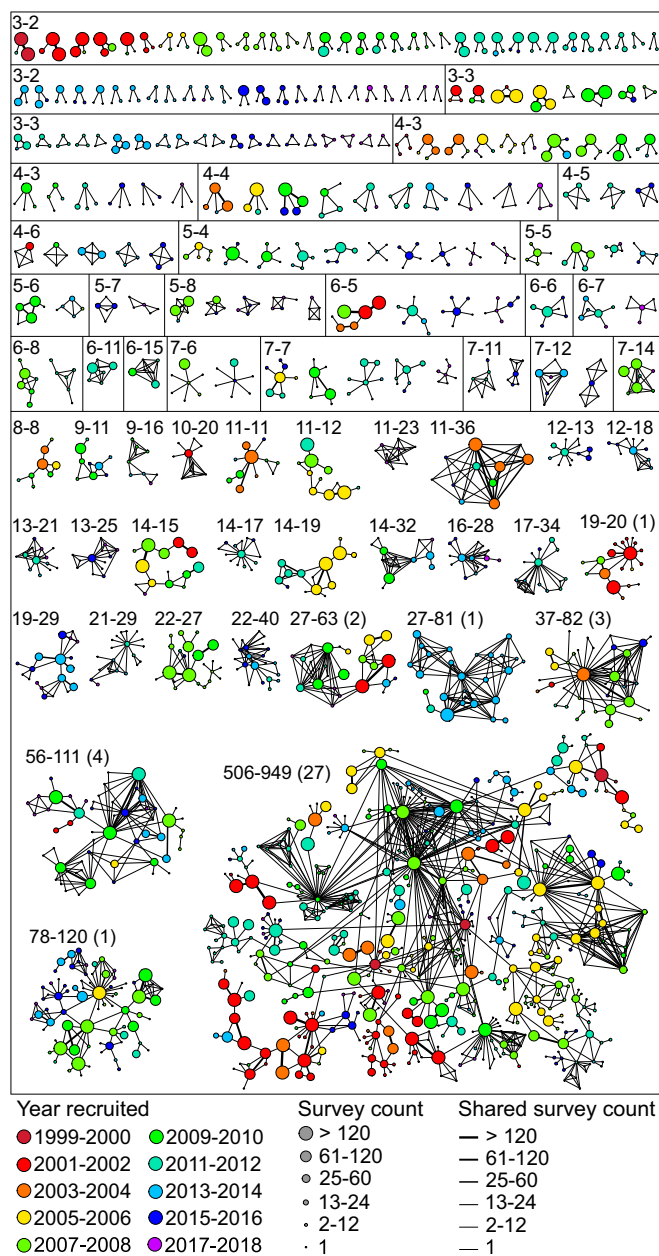
trade-off between retention and accuracy further by categorizing participants by demographic and interest factors cited by others as potentially influential (19, 23–25). Initially, the highest retention rates (>90% participation beyond the first survey) were associated with participants who had previously been involved in other citizen science projects, already had some expertise in identifying birds, and were older than 40 y of age when they joined (*SI Appendix*, Fig. S4). One year later, these patterns had become more pronounced. After 5 y, age and prior bird identification skill were the only influential demographic characteristics affecting retention. The highest levels of accuracy were achieved by younger participants, followed by middle-aged (40–60 y) participants and those with birding expertise (*SI Appendix*, Fig. S4). Gender was the least influential demographic characteristic in retention or accuracy, although males were marginally more persistent than females as a function of time (one survey, 0.9%; 1 y, 2.5%; 5 y, 3.6%; *SI Appendix*, Fig. S4).

Fig. 3 demonstrates the trade-offs between rates of retention and accuracy, as participant subpopulations move from their initial status (after one survey and one bird) to their "seasoned" status 1 y later, or two snapshots along the trade-off frontier (i.e., Fig. 2*C*). Across all demographic groups, initial spread in retention (81–94%, Δ = 12.3) was approximately half of what it reached at 1 y of surveying (43–64%, Δ = 21), whereas the pattern in accuracy was the opposite (initial Δ = 15%; seasoned Δ = 8.7%). Thus, attempting to manipulate the target audience as a function of demographic characteristics should have a larger influence on retention than on accuracy. Some demographic characteristics consistently predicted better performance and higher retention (e.g., involvement in other citizen science projects, birding experience) whereas others displayed trade-offs (e.g., age). For instance, older individuals (>60 y) began as the least accurate age class and had become the most accurate by 1 y, although all age classes improved (Fig. 3).

A profiling approach (e.g., ref. 32) would suggest that favoring older individuals (>40 y) with at least intermediate experience in identifying birds who also had previous citizen science experience would deliver the highest levels of retention and accuracy at 1 y. This approach would also suggest that overall project statistics could be improved by avoiding younger individuals lacking bird or citizen science experience, even though younger individuals will eventually become the most accurate performers (*SI Appendix*, Fig. S4*G*). In fact, the single largest demographic subpopulation in COASST (7.6%) is older (>60 y) females with previous citizen science experience and little to no birding expertise: neither the most persistent or accurate group nor the least. We suggest that in-depth knowledge of participant demographic characteristics is a double-edged sword. Knowing the relative performance of each demographic- and/or interest-based subpopulation can be used to tune project design toward favoring best-performing individuals (i.e., high grading) or toward providing more interaction and project support to increase the performance of all classes to project-acceptable levels.

Science as a Social Enterprise. To understand the distribution and influence of social relationships on project statistics, we categorized participants surveying for ≥6 mo (*n* = 1,815) according to their expressed preferences as to the number and identity of survey partners. Individuals joining COASST are asked to survey in pairs, and the majority of participants (51%) did so, although many had more than one partner over time (Figs. 4 and 5). "Loners" were the second most prevalent social category (24%), and 36% of those surveyed exclusively alone. We classified the remaining 25% of participants as "gregarious" (Fig. 4). By and large, these individuals also surveyed as pairs, but with many more unique partners. When categorized dichotomously as loner or social, being social had no apparent influence on participant

Fig. 3. Project-level trade-offs extracted from subpopulation-specific learning and retention curves (e.g., *SI Appendix*, Fig. S4) as the mean and 95% CI of accuracy and retention, respectively, at two points on each curve. Initial (upper cluster) indicates the percentage of the population that perform more than one survey and the accuracy at correct species identification modeled at the first bird encountered; seasoned (lower cluster) indicates the percentage of the population persisting for >1 y and the accuracy at the 30th bird encountered (i.e., a typical year's accumulation of carcasses in the Pacific Northwest outer coast, assuming monthly surveys of a single beach). Black symbols denote all participants (i.e., project-wide statistics). All other colors denote inclusive classes for which demographic characteristics and interests are known.

retention or accuracy (*SI Appendix*, Fig. S4 *E* and *J*). Thus, sociality appears to have no overt influence on program statistics.

However, when COASST is examined as a social network, whereby individuals are connected to other individuals with whom they have surveyed, a different picture emerges. The majority of participants belonged to networks of ≥4 people, with the largest network exceeding 500 individuals (Fig. 5 and *SI Appendix*, Fig. S5). These people are not all known to each other; in COASST, social networks are built over time. That is, participants recruited guests, some of whom become regular partners, with some of those eventually starting their own survey sites and inviting others to participate. Just more than 2% of participants were classified as "nexus" people, a subset of gregarious people defined by having >10 unique partners (Fig. 4). By definition, nexus people were the most well-connected, falling principally in the largest networks (Fig. 5 and *SI Appendix*, Fig. S5). On average, nexus people each recruited nine other individuals (maximum, *n* = 30), compared with loners and pairs (averaging <0.5 recruits per participant) and other gregarious individuals (2 recruits). Nexus people were not necessarily enthusiasts, in the sense of conducting the vast majority of the work themselves (sensu; ref. 10), nor were they extraordinary performers, in the sense of hobbyists (sensu; ref. 19) or top players (sensu; ref. 33). Demographically, they largely mirrored the overall COASST population, with the exception of significantly fewer young individuals compared with the wider COASST population (5.4% observed vs. 19.6% expected; *P* = 0.024). In sum, the additional value of nexus participants and, to a lesser extent, all gregarious individuals was in their ability to bring new people into the project.

Discussion

It's About Design. Our results extend the work of others suggesting that the pattern of contributor effort in crowdsourced, online citizen science projects follows a Pareto curve (8, 33, 34) by revealing that this "peak" in the active citizen science retention landscape is only one of several locations (Fig. 1*B*). Translated into participant numbers, even though online projects may recruit thousands of participants annually, they may retain only a few hundred, whereas a high-retention, hands-on project may need to recruit only 100 participants because almost all will be retained. However, successful, stable hands-on projects ran the gamut of retention, from extremely high-retention projects to those in which the majority of contributors were dabblers. We suggest that the retention landscape emerges from a variety of project designs beyond the online vs. hands-on division, including but not limited to blitzes, opportunistic surveys, and standardized, regular monitoring. Fleshing out the landscape with many more projects and linking retention to performance at individual and project levels will facilitate a more comprehensive and systematic examination of the intersections among project design, optimization of scientific achievement, and benefits to the participating individual.

Studies of motivation in environmental volunteerism have long suggested that certain activities recruit individuals who come to the project with a high level of expertise developed over a lifetime of practice. Although these amateur experts may contribute to a citizen science project out of a desire to assist in and/or be part of science (24), they also clearly regard content-relevant projects as providing an opportunity to extend their hobby (35). As a result, projects designed to take advantage of outdoor national pastimes may attract a ready supply of near-professionals who, finding their functional requirements fulfilled (22), maintain their involvement for years. Given the dramatic shifts in human populations toward urban environments (36) and online activities (37), it is an open question whether these classic naturalists will persist in numbers sufficient to fuel this end of the retention landscape continuum. Alternately, emergent shifts in society toward online activity may be producing future generations

Fig. 4. Social dimensions within COASST. Mean number of partners per survey as a function of the total number of unique partners across all surveys for each participant (including guests; *n* = 3,286), with each participant shown as a dot. Dot size is scaled to cumulative surveys performed; color saturation is scaled to number of individuals within a particular mean and cumulative partner intersection. Dashed lines delineate social categories: loners (*Bottom Left*), pairs (directly above loners), and gregarious (all other participants). Nexus participants (shaded in red) are a special class of gregarious individuals defined by >10 unique partners and >60 surveys.

Fig. 5. The COASST social network. Each node (circle) represents an individual coded by start date (color) and number of surveys performed (size), and each link (connecting line) represents surveys in common coded by total count (line thickness). Numbers indicate the number of people (nodes), survey partnerships (links), and nexus people, respectively, in each network grouping. Loners and pairs are excluded.

themselves to highly attractive institutions that are also tightly rooted in place, including parks, zoos and aquariums, and museums. For instance, 300 million people annually visit US national parks, and therefore park-centered blitz projects may act as an entrée to further public involvement in science (5). Our results from the COASST project indicate that prior participation in citizen science had a significantly positive impact on retention (Fig. 3 and *SI Appendix*, Fig. S4), hinting that a "career progression" of project participation (sensu 8) could start with dabbling and move toward more sustained involvement.

Hands-on projects between these poles demonstrate perhaps the best potential to realize scientific outcomes on the part of project organizers and personal learning and other types of functional fulfillment on the part of the individual participants (2), as contributors have the opportunity to practice skills learned within the context of the project while also gaining understanding and awareness of scientific pattern, process, and function beyond what any single individual could experience (31). It is also in this space that the trade-offs between retention and task performance at the project level—or persistence and learning seen from the participant's perspective—emerge most strongly (Figs. 2 and 3 and *SI Appendix*, Fig. S4). Although individual learning curves are no doubt project-specific, retention/accuracy trade-offs are, by definition, a general feature of hands-on deductive data collection projects. We suggest that knowing the shape of the trade-off frontier (e.g., Fig. 2*C*) between these two basic metrics of citizen science, both over all participants and by sub-population, can be invaluable in focusing design effort where it is needed (i.e., groups with lower performance and/or shorter retention times) and, more broadly, in understanding how the project effectively favors some individuals over others, at the very least via differential retention. Clearly, more programs need to delve into the relationships among demographic characteristics, interests, and program statistics, including retention and metrics of participant accuracy directly related to data quality.

Can Learning Happen at All Participant Levels? If learning is truly lifelong, do all participants in citizen science reap this benefit? In COASST, even self-rated experts posted accuracy gains over their first year (Fig. 3 and *SI Appendix*, Fig. S4), a finding echoing online (8) and hands-on (41) studies. At the other end of the spectrum, the question remains open whether dabblers are able to experience measurable learning gains.

Conceptually, low-dabbler, high-retention projects should afford an opportunity to push the majority of the participant corps up the learning curve, as most individuals conduct tasks repeatedly. Continuing participants can certainly assimilate content knowledge and associated skills specific to a project (15), and some are able to demonstrate practices of science in the absence of prompting (42, 43). Although several studies have examined participant proficiency in hands-on citizen science projects, usually relative to an expert population (e.g., ref. 44), relatively few have examined shifts in data collection accuracy through time, or the learning curve (e.g., ref. 41). Our study is one of the first to report measured performance over time in the participant population, and our results suggest that all participants who persisted past one survey gained in their ability to correctly identify carcasses to species regardless of their starting point or demographic characteristics (Fig. 3).

However, it is also the case that participants started, on average, at 78% accuracy (Fig. 2*B*), a rate suggesting that they were already experts. In fact, this is exactly the case, albeit not because carcass identification has been learned elsewhere, as the majority of participants (51%) self-rated as novice birders at joining (31). Furthermore, the sheer number of species recorded by the program (>150 species) belies the hypothesis of prior expertise.

Year recruited

- 🔴 1999–2000
- 🔴 2001–2002
- 🟠 2003–2004
- 🟡 2005–2006
- 🟢 2007–2008
- 🟢 2009–2010
- 🟢 2011–2012
- 🔵 2013–2014
- 🔵 2015–2016
- 🟣 2017–2018

Survey count

- ● > 120
- ● 61–120
- ● 25–60
- · 13–24
- · 2–12
- · 1

Shared survey count

- — > 120
- — 61–120
- — 25–60
- — 13–24
- — 2–12
- — 1

of cyber-savvy individuals who are poised to contribute to the emergent set of online citizen science opportunities (38), including the burgeoning number of projects focused on environmental and conservation science (39).

The other end of the hands-on retention continuum appears to be projects intentionally focused on one-off involvement, exemplified by blitzes (29), in which a large number of individuals are recruited to come together and simultaneously engage in a task. Although blitzes may be characterized by high rates of dabbling and low retention, these programs can return significant scientific results when targeted toward specific issues (e.g., range extension of invasive species) (40). Because hands-on blitz projects and other one-off events are spatially centralized, they lend

Rather, we assert that this level of performance at taxonomic identification is the effect of training (28) and project design. Breaking the scientific process down into multiple simple tasks is a successful strategy for returning high-quality results (8). In COASST, the process of taxonomic identification has been effectively disarticulated into the collection of evidence and comparative analysis using these data, resulting in participant deductions that are based on fact rather than hunch (28). This suggests that citizen science participants can quickly become competent such that high-quality datasets can be reliably produced even by relatively novice nonexperts if the underlying science process is explicitly, and project-specifically, revealed.

Citizen Science for Everyone. As science seeks to truly open the doors to participation, citizen science holds an additional promise as a free-choice learning experience contributing to the public's understanding of science and the scientific process (45). Individuals are attracted to projects and remain as long-time contributors because their interests and motivations match those of the project (20). In this idealized worldview, everyone has the potential to be a part of the science team, contributing to scientific outcomes that are meaningful to them and to science (46).

To date, efforts to understand the diversity of participants in citizen science have principally focused on relative contributor effort and abilities, touching only lightly on demographic disparities and social dimensions. Are there "optimal personas" in citizen science? Certainly "enthusiasts" and "top players" may be people with particular skills and unique ways of thinking beneficial to project success (10, 33). Although the COASST results point to some demographic combinations as persistent and highly accurate data collectors (Fig. 3), a combination certainly advantageous at the project level, they also indicate both trade-offs and synonymies over time (*SI Appendix*, Fig. S4). For instance, older participants started as the age class least accurate at identifying carcasses to species, became the most accurate within 1 y, and then lost that status again to younger individuals by the 5-y mark. These results suggest that a profiling approach may be difficult, if not fruitless, for projects retaining individuals over years. Finally, our finding that the most socially influential people in terms of recruiting others into the project (i.e., nexus people; Figs. 4 and 5) were not necessarily the highest-performing individuals suggests that an optimal persona—long-lasting, highly accurate at data collection, and highly efficient at recruiting others to sustain the program—may be a myth.

Pandya (21) points to the lack of underrepresented and underserved peoples in citizen science. Cooper and Smith (23) suggest that the inclusion, and exclusion, of women as citizen science participants may be tied to the designed sense of membership (female-dominant) vs. competitiveness (male-dominant) across projects. Ganzevoort et al. (24) suggest that out-of-doors, nature-based citizen science appeals predominantly to older individuals. Clearly, there are reasons why specific instantiations of citizen science may drive people away from the opportunity rather than toward it. Understanding motivations to join, as well as not to join, as a function of participant identity is a fundamental part of intentional project design.

Although each project will have a different spread of retention and task performance across the participating population, our results suggest that a profiling approach might increase data collection accuracy in COASST by <10 percentage points but depress participation to potentially unsustainably low numbers. Conversely, attention to the factors resulting in the relatively lower rates of retention and/or accuracy of younger participants, and those without previous citizen science experience or bird identification expertise (e.g., Fig. 3, lower left of the 1-y cluster) could significantly improve project statistics and might increase project attractiveness to a wider population base. Thus, although we are in favor of a demographic approach to the study of retention and performance, we argue that its use for inclusive design is statistically prudent and socially responsible.

Conclusions: Serendipity and Emergence. Citizen science across the continuum from online to hands-on involves millions of people in discovery and actionable science (*SI Appendix*, Fig. S1). This burgeoning science team creates an opportunity for serendipity and emergence. The sudden discovery of a new star formation, or a new species, or a new protein structure: these saltatory advancements appear to take actions of the many to produce. And the actions of the many also allow realization of patterns in space and time unknowable by the individual (27): shifts in the timing of spring flowering, changes in migratory pathways, continental spread of a disease, or climate-induced massive mortality. Such "big science" requires big data. The COASST case study clearly indicates that persistent nonexperts can become highly accurate at making complex deductions (i.e., species identification) from directly collected evidence (e.g., foot type, measurements). Additional studies on the patterns of, and trade-offs between, participant retention and data collection accuracy in hands-on citizen science are essential to address the skeptical among the mainstream scientific community (12). We suggest that high-quality data can be successfully produced by appropriately designed citizen science projects. We envision a future in which all of the public are included in these efforts, and we pose the question: what advancements might we collectively realize?

Materials and Methods

Data on 3,286 COASST participants (1999–2018) were obtained from the COASST database. Demographic and interest information are voluntarily provided by participants upon joining with their informed consent for future use under University of Washington Institutional Review Board (IRB) protocols 37516 and 47963. We defined regular participants (n = 2,356) as those who attended an expert-led COASST training and/or who were trained on the beach by others and subsequently began monthly surveys. Guests (n = 930) were defined as individuals recruited by regular participants to fill in as occasional partners, but who did not survey regularly, logging no more than five surveys. COASST staff, interns, agency partners, and individuals conducting surveys as part of their job (e.g., tribal biologists, National Park Service Rangers) were excluded from this analysis.

We coded the social status of all regular participants who surveyed for at least 6 mo (n = 1,815) as follows: loners, ≥50% of their surveys alone and <5 unique partners; and pairs, <50% of their surveys alone and <25% of their surveys with multiple partners, with <5 unique partners. All remaining participants were termed gregarious. Within that category, nexus individuals were defined as those participants with >10 unique partners and >60 surveys.

Details on the measures used and statistical models constructed are provided in *SI Appendix, Supplementary Text*. Unless otherwise stated, all analyses were carried out by using R version 3.3.4 (47). All COASST data not protected by IRB regulation are available in summarized form on our website (www.COASST.org), and are available in raw form on request and with a signed data sharing agreement.

1. Bonney R, et al. (2014) Citizen science. Next steps for citizen science. *Science* 343: 1436–1437.
2. Dickinson JL, et al. (2012) The current state of citizen science as a tool for ecological research and public engagement. *Front Ecol Environ* 10:291–297.
3. McKinley DC, et al. (2017) Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol Conserv* 208:15–28.
4. Bonney R, et al. (2009) Citizen science: A developing tool for expanding science knowledge and scientific literacy. *Bioscience* 59:977–984.
5. Watkins T, Miller-Rushing AJ, Nelson SJ (2018) Science in places of grandeur: Communication and engagement in national parks. *Integr Comp Biol* 58: 67–76.
6. Trouille L, Lintott C, Miller G, Spiers H (2017) DIY Zooniverse citizen science project: Engaging the public with you museum's collections and data. *MW17: MW 2017*. Available at https://mw17.mwconf.org/paper/diy-your-own-zooniverse-citizen-science-project-engaging-the-public-with-your-museums-collections-and-data/. Accessed November 28, 2018.
7. Follett R, Strezov V (2015) An analysis of citizen science based research: Usage and publication patterns. *PLoS One* 10:e0143687.
8. Sauermann H, Franzoni C (2015) Crowd science user contribution patterns and their implications. *Proc Natl Acad Sci USA* 112:679–684.
9. Theobald EJ, et al. (2015) Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biol Conserv* 181:236–244.
10. Boakes EH, et al. (2016) Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. *Sci Rep* 6:33051.
11. Kosmala M, Wiggins A, Swanson A, Simmons B (2016) Assessing data quality in citizen science. *Front Ecol Environ* 14:551–560.
12. Burgess HK, et al. (2017) The science of citizen science: Exploring barriers to use as a primary research tool. *Biol Conserv* 208:113–120.
13. Gallo T, Waitt D (2011) Creating a successful citizen science model to detect and report invasive species. *Bioscience* 61:459–465.
14. van der Wal R, Sharma N, Mellish C, Robinson A, Siddharthan A (2016) The role of automated feedback in training and retaining biological recorders for citizen science. *Conserv Biol* 30:550–561.
15. Masters K, et al. (2016) Science learning via participation in online citizen science. *J Sci Comm* 15:A07.
16. Swanson A, Kosmala M, Lintott C, Packer C (2016) A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conserv Biol* 30:520–531.
17. Cooper S, et al. (2010) Predicting protein structures with a multiplayer online game. *Nature* 466:756–760.
18. Dickinson JL, Zuckerberg B, Bonter DN (2010) Citizen science as an ecological research tool: Challenges and benefits. *Annu Rev Ecol Evol Syst* 41:149–172.
19. Jones MG, Corin EN, Andre T, Childers GM, Stevens V (2017) Factors contributing to lifelong science learning: Amateur astronomers and birders. *J Res Sci Teach* 54: 412–433.
20. West S, Pateman R (2016) Recruiting and retaining participants in citizen science: What can be learned from the volunteering literature? *Citiz Sci Theor Prac* 1:A15.
21. Pandya R (2012) A framework for engaging diverse communities in citizen science in the US. *Front Ecol Environ* 10:314–317.
22. Clary EG, Snyder M (1999) The motivations to volunteer: Theoretical and practical considerations. *Curr Dir Psychol Sci* 8:156–159.
23. Cooper CB, Smith JA (2010) Gender patterns in bird-related recreation in the USA and UK. *Ecol Soc* 15:A26268198.
24. Ganzevoort W, van den Born RJ, Halffman W, Turnhout S (2017) Sharing biodiversity data: Citizen scientists' concerns and motivations. *Biodivers Conserv* 26:2821–2837.
25. Asah ST, Blahna DJ (2013) Practical implications of understanding the influence of motivations on commitment to voluntary urban conservation stewardship. *Conserv Biol* 27:866–875.
26. Parrish JK, et al. (2018) Exposing the science in citizen science: Fitness to purpose and intentional design. *Integr Comp Biol* 58:150–160.
27. Woolley JP, et al. (2016) Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives. *BMC Med Ethics* 17:33.
28. Parrish JK, et al. (2017) Defining the baseline and tracking change in seabird populations: The Coastal Observation and Seabird Survey Team (COASST). *Citizen Science for Coastal and Marine Conservation*, eds Cigliano JA, Ballard HL (Routledge, New York), pp 19–38.
29. Lundmark C (2003) BioBlitz: Getting into backyard biodiversity. *Bioscience* 53:329.
30. Roos S, Johnston A, Noble D (2012) UK hedgehog datasets and their potential for long-term monitoring. BTO Research Report 598 (British Trust for Ornithology, Thetford, UK).
31. Haywood BK, Parrish JK, Dolliver J (2016) Place-based and data-rich citizen science as a precursor for conservation action. *Conserv Biol* 30:476–486.
32. Hunter J, Alabri A, van Ingen C (2013) Assessing the quality and trustworthiness of citizen science data. *Concurr Comput Prac Exper* 25:454–466.
33. Khatib F, et al. (2011) Algorithm discovery by protein folding game players. *Proc Natl Acad Sci USA* 108:18949–18953.
34. Segal A, et al. (2015) Improving productivity in citizen science through controlled intervention. *Proceedings of the 24th International Conference on World Wide Web, Florence, Italy. May 18–22, 2015* (ACM, New York), pp 331–337.
35. Rotman D, et al. (2014) Motivations affecting initial and long-term participation in citizen science projects in three countries. *iConference 2014 Proceedings, Berlin, Germany. March 4–7, 2014* (iSchool, Grandville, MI), pp 110–124.
36. United Nations, Department of Economic and Social Affairs, Population Division (2018) World Urbanization Prospects: The 2018 Revision (UN DESA, New York).
37. Ryan C, Lewis JM (2017) Computer and internet use in the United States: 2015. American Community Survey Reports ACS-37 (US Census Bureau, Washington, DC).
38. Newman G, et al. (2012) The future of citizen science: Emerging technologies and shifting paradigms. *Front Ecol Environ* 10:298–304.
39. Pocock MJO, Tweddle JC, Savage J, Robinson LD, Roy HE (2017) The diversity and evolution of ecological and environmental citizen science. *PLoS One* 12:e0172579.
40. Cohen CS, McCann L, Davis T, Shaw L, Ruiz G (2011) Discovery and significance of the colonial tunicate *Didemnum vexillum* in Alaska. *Aquat Invasions* 6:263–271.
41. Kelling S, et al. (2015) Can observation skills of citizen scientists be estimated using species accumulation curves? *PLoS One* 10:e0139600.
42. Trumbull DJ, Bonney R, Bascom D, Cabral A (2000) Thinking scientifically during participation in a citizen-science project. *Sci Educ* 84:265–275.
43. Straub M (2016) Giving citizen scientists a chance: A study of volunteer-led scientific discovery. *Citiz Sci Theor Prac* 1:A5.
44. Crall AW, et al. (2011) Assessing citizen science data quality: An invasive species case study. *Conserv Lett* 4:433–442.
45. Bonney R, Phillips TB, Ballard HL, Enck JW (2016) Can citizen science enhance public understanding of science? *Public Underst Sci* 25:2–16.
46. Haywood BK, Besley JC (2014) Education, outreach, and inclusive engagement: Towards integrated indicators of successful program outcomes in participatory science. *Public Underst Sci* 23:92–106.
47. R Core Team (2017) R: A Language and Environment for Statistical Computing, version 3.3.4 (R Foundation for Statistical Computing, Vienna).