Ideas for Citizen Science in Astronomy

Phil Marshall,[1,2] Leigh Fletcher,[2] and Chris Lintott[2]

[1] *Kavli Institute for Particle Astrophysics and Cosmology, P.O. Box 20450, MS 29, Stanford, CA 94309, USA.*

[2] *Department of Physics, Denys Wilkinson Building, University of Oxford, Keble Road, Oxford, OX1 3RH, UK.*

Abstract

We review the relatively new, internet-enabled, and rapidly evolving field of citizen science, focusing on projects in astronomy and solar system science. In doing so, we look for ideas from which astronomy could benefit in the future. We consider contributions to science in the form of observations, data processing, data modeling and the design of new scientific inquiries, and from this experience derive some commonalities between the most successful projects, and use them to suggest some guidelines for future projects. The limits of citizen science are not yet well understood, but we make some speculations for citizen astronomy in particular.

CONTENTS

1 Introduction (2 pages)

The term "Citizen Science" refers to the activities of people who are not paid to carry out scientific research, but nevertheless make intellectual contributions to scientific research in their spare time. These contributions are diverse, both in type and research area. The people who make those contributions may come from all walks of life [REF?]. This review is about the science projects they have participated in to date, the tasks they have performed, and how astronomy has benefited – and could benefit further – from their efforts.

Citizen involvement in science pre-dates the profession itself, and there is a long and honourable tradition of amateur observers making important discoveries and significant sustained contributions. However, the advent of the world wide web has changed the face of professional and amateur collaboration, providing new opportunities and accelerating the sharing of information. People are now connected to each other in a way that has never happened before. Professional scientists can interact with citizens via a range of web-based media, including purpose-built citizen science websites which increase the potential for shared data analysis and exploration as well as data collection. Meanwhile, communities of citizens have sprung into existence as like-minded people have been able to find and talk to each other in a way that is almost independent of their geographical location. The result has been an exponential increase in citizen involvement in science. The field is evolving very quickly, with more and more professional scientists becoming aware of the possibilities offered by collaborating with, for example, specialists operating outside the usual parameters of professional astronomical observation, or tens of thousands of people eager to perform microtasks in their spare time. Our aim in this work is to review the scientific literature as it stands for ideas implemented in citizen science projects, primarily in astronomy but occasionally in other fields, and then produce a summary of successful project characteristics for future research groups to learn from.

As our title states, this is a review of ideas for astronomy. We will look forward as well as back, and try to answer the questions: How can the full potential of citizen science be realised in astronomy? What are the particular niches that citizen science can fill, in our field? How might it contribute to the solutions of the Big Data problem in astronomy?

This review is organised as follows. We survey the contributions to science that citizens have made to date, organized according to the stage of the scientific enquiry that those contributions fell into. Astronomy research typically starts with observations: so do we, in Section 2. We then proceed through a very brief discussion of citizen instrumentation, and then consider data processing, data modeling and finally citizen-led enquiry in Sections 2.3–5. With this overview in place, we review in Section 6 the literature on, and the collected experience of, the population of citizens who have taken part, or are currently taking part, in scientific research. We then turn to the future, and speculate on how citizens might contribute to astronomy there (Section 7); in Section 8 we consider possible limits to citizen science, including challenges associated with data rates and volumes, data complexity, the difficulties of large-scale collaboration, and the barriers to accessibility. Finally, we make some concluding remarks in Section 9, and suggest some guidelines for the implementation of future citizen science projects in astronomy.

2 Data Acquisition: Citizen Observing (5 pages)

There is currently an active community of well-equipped amateur observers making astronomical observations of great utility. There are also many other citizens observing the night sky with less sophisticated equipment, and/or less enthusiasm – and as we shall see, there are plenty of citizens making astronomical observations almost inadvertently. What astronomical data are the citizenry taking, and what is it being used for?

2.1 Active Observing (3 pages)

The steady improvements and increasing affordability of digital technology, in addition to the ease of data sharing and communications, have considerably expanded the realm of amateur astronomy in the past two decades. Observers have always been passionate about their pastime, from painstakingly recorded hand drawings of the planets and their positions in the night sky, to hours devoted to the monitoring and tracking of astronomical events. There is a long and proud history of contributions from ?amateurs? to the field, driven by a desire to share the results of their hobby with others. But it has only been in the past few decades that regular and systematic pipelines have been in place to share data between amateurs and professionals, enticing passionate observers to shape their observations to address scientific questions. In this section, we review some of the citizen contributions to active observations of the night sky. Passive contributions, via data mining and web-based studies, will be described in subsection 2.2 below.

Why is active citizen observing beneficial to professional astronomers? The key advantage is time ? the very core of astronomy is observational, and intense competition for resources means that professional astronomers are unable to continuously monitor all the objects of interest in the night sky. Indeed, professional observatories are always oversubscribed, with resources concentrated on one area of sky, or one astronomical question. Observations from professional telescopes, space-based observatories and visiting planetary missions are often sparse and poorly sampled due to extreme competition for resources. Such observations are rarely tuned to the optimum timescales for scientific enquiry ? for example, determinations of meteor frequencies on short timescales (minutes), or slow evolution of giant planets on longer timescales (years and decades). Amateur observations can be frequent and repetitive to provide monitoring of astronomical targets, and are naturally well sampled across the globe during an exciting event of interest.

The second, related, advantage is that of flexibility ? whenever a new phenomenon is discovered (e.g., a new comet, or anything changing the appearance of the familiar planetary discs), observers will be keen to catch a glimpse irrespective of the scientific value of their observations. This reaction can be near instantaneous, compared to the need to allocate telescope resources among the professional community. The third benefit is contextual ? global maps of an object's visible albedo provide useful additional constraints on a process of interest, especially when near-simultaneous professional observations occur in a vastly different wavelength range (e.g., UV or infrared); over a narrow spatial region (e.g., close-in studies by spacecraft); or employ spectroscopic techniques that don?t yield images. Comparisons with the visible albedo provided by citizen scientists can therefore play an extremely useful role in planetary missions, both for context and for observation planning. Active observation by enthusiastic citizens contributing directly to scientific progress, combined with the excellent communication channels that have opened up in the past two decades, and the digital technology that allows astrophotographers to shine, is beginning to plug this observational gap.

As a result of these benefits, a few professionals are very closely linked with the amateur community, organising publications and assisting in observation planning to maximise science return. New planetary missions frequently involve scientists to serve as an interface with the amateur community (this is particularly true to the 2016-17 Juno mission to Jupiter). In return, amateurs trigger alerts through online forums or emails to professional networks, announcing a tentative new discovery to the world for scrutiny and follow-up. The benefits of strong professional-amateur collaboration will be repeated in the following examples where active observations have contributed directly to our understanding of our place in the universe. We do not intend a complete review of all examples of Pro-Am collaboration, but a few representative examples. For a detailed review of the methodology employed by amateur observers, the reader is referred to **?**). Within our solar system, active observations fall into two interconnected categories: Discovery and Monitoring, where each will almost certainly lead to the other.

2.1.1 Discovery-Class Science Discovery-class citizen science is enabled by the quantity of data obtained and the extreme familiarity of citizen astronomers with a particular region, planet or nebula, allowing them to immediately identify peculiarities or new features (e.g., meteorological activity on giant planets, **?**). Solar system objects moving against the fixed-star background can be detected in a set of CCD frames either by eye or by automated software. The position of this new object is then compared to existing catalogues, and if no existing details are found then the new discovery and its ephemerides can be reported to the IAU Minor Planet Centre[1]. If observations are repeated for at least two nights by one or several observers, then a new denomination is provisionally assigned to the discovery. An electronic circular then reports the discovery to the wider world. Suitable targets for these searches include near-earth asteroids (NEAs, with orbits intersecting those of the terrestrial planets), main belt asteroids between Mars and Jupiter, and comets making their journey towards the Sun from the outer solar system.

The recent close flyby of asteroid 2012 DA14 on February 15th 2013 was initially reported by a team of amateur observers affiliated with the La Sagra Sky Survey at the Astronomical Observatory of Mallorca. Amateur observers still contribute to the discovery and photometric imaging of comets, and many amateurs have managed to become associated with new cometary discoveries, including David Levy, part of the team that discovered Shoemaker-Levy 9 before its spectacular impact with Jupiter. As with asteroids, the majority of new comet discoveries are made by automated surveys, but a small and stable number of discoveries come from amateurs with small telescopes, typically in regions poorly covered by survey telescopes (e.g., regions close to the Sun). C/2011 W3 Lovejoy, a Kreutz sungrazer comet, is one such example. C/2012 S1 ISON was spotted by A. Novichonok and V. Gerke in images from the International Scientific Optical Network. One might imagine that, as observing technology improves, citizen discoveries via active observation might extend out to the distant Kuiper Belt, although many of these targets are so dim that they are unavailable to all but the largest of apertures.

[1] http://www.minorplanetcenter.net

Although survey telescopes provide the vast majority of modern discoveries, citizen observations allow detailed characterisation of physical and orbital characteristics of these newly discovered solar system bodies, and amateur-led contributions are published online[2]. There is a need to continuously track Near Earth Asteroids as gravitational effects can significantly alter orbital trajectories over time, and this astrometry for both asteroids and comets can be provided to the Minor Planet Centre for collation and use in computing orbital trajectories. Amateur monitoring of a comet?s coma, dust and plasma tails and their photometric parameters can reveal dynamic structures and determine the locations of active venting regions. Photometric monitoring of comets over time of the magnitude and tail characteristics provides insights into the levels of activity as the comet moves along its trajectory, and can also reveal outbursts and other events associated with the outgassing. Photometric monitoring of an asteroid as it rotates provides information on the physical parameters, such as the shape, rotation rate and orientation. Observationas of mutual encounters can also yield information on the density, and hence the composition of the asteroids.

Amateurs are also contributing to the search for a sub-category of objects with a detectable cometary coma within the asteroid belt. Recent discoveries of these Main Belt Comets, which appear to be asteroids that are actively venting their volatiles at perihelion, are beginning to blur the distinction between asteroids and comets. The T3 project, a collaboration between the University of Rome and several amateur observers, began in 2005 with the detection of a coma around asteroid 2005 SB216 (**?**), and has gone on to detect at least eight main belt comets (**?**). Comae may be detected as an extended FWHM of the asteroid compared to

[2] http://www.minorplanet.info/mpbdownloads.html

the background stars, or visually distinctive cometary comae. Recent discoveries of cometary activity can be found here: http://schiaparelli204.wordpress.com/

Beyond our solar system, amateurs have contributed to exoplanetary transit discoveries, attempting to measure the 1planet transits in front of its parent star. **?**) points out three methods where amateurs can contribute to characterising exoplanetary systems ? (i) by frequent observations of known transits to refine ephemeris; (ii) searching for transit time variations that can reveal additional planets in a system; and (iii) searching for previously unidentified transits in known planetary systems (e.g., the discovery of the transit of HD 80606b from a 30 cm telescope near London, **?**).

2.1.2 Long Baseline Monitoring The diverse and dynamic environments of our solar system are tantalising targets for amateur observations, both due to their ease of identification and the chance that some unusual, never-before-seen phenomenon might present itself before your eyes. Long-term quasi-continuous monitoring of these objects is highly desirable, but impossible to achieve from over-burdened professional telescopes. Citizen science is able to plug that gap, providing an expanding dataset of observations to trace the day-to-day evolution of the inner and outer planets. Jupiter?s size and colourful cloud contrasts, with its striped appearance and large vortices, make it an ideal first target for budding astronomers. Winds, waves, storms and plumes shape the day-to-day appearance, with cloud structures being blown east and west by powerful zonal jets. The jet velocities, and the dispersion of cloud material, reveals processes at work beneath the visible ammonia ice clouds, and serve as a probe of the fluid dynamics at work. Tracking the motion of features requires correlation of images obtained night after night; with filters sounding in and out of strong jovian methane absorption features providing a three dimensional view of the planet?s cloud decks. Both colour composites and raw filtered images are uploaded to online servers, organised by date and time, such as the Planetary Virtual Observatory and Laboratory (PVOL, http://www.pvol.ehu.es/pvol) maintained for the International Outer Planets Watch (IOPW) (**?**). The global distribution of giant planet observers permits global monitoring of Jupiter and Saturn as they rotate over 10 hours. Descriptive records of morphological changes and events are maintained and continuously updated by organisations such as the British Astronomical Association (BAA) and Association of Lunar and Planetary Observers (ALPO). Those images can be used by amateurs and professionals alike to quantitatively study the zoology of activity; from measuring wind speeds (**?**); investigating the strength and changes to the large vortices (e.g., the 2006 reddening of Oval BA, **?**); to determining the life cycle of the belt/zone structure (**??**).

Similar monitoring studies are underway for Saturn, Uranus and Neptune, with increasing levels of difficulty. Saturn's appearance is typically more subdued than that of Jupiter, but 20-30 cm diameter telescopes are capable of resolving small convective cloud activity. A close collaboration between amateurs and Cassini spacecraft scientists allows correlation of lightning-related radio emissions detected by the spacecraft with visible cloud structures on the disc (known as Saturn Storm Watch) (e.g., **?**), which would not be possible with the targeted regional views provided by Cassini's cameras alone. This provides insights into the moist convective processes thought to power the dynamics of the giant planet, and is a good example of how citizen science can support an international planetary mission. Amateur observations of Uranus and Neptune are in their infancy and require telescopes with diameters exceeding 25 cm, but there have been confirmed reports of atmospheric banding and discrete cloud features when near-infrared filters are used to maximise contrast between white clouds and the background and long exposure times of tens of minutes.

Active citizen observing also provides long-term monitoring in the inner solar system. Venus? photochemical smog shields the planet's surface from view, but discrete cloud features can be used to study the super-rotation of the Venusian atmosphere and the occurrence of a mysterious ultraviolet absorber high in the planet?s atmosphere (i.e., using near-UV filters). The Venus Ground-Based Image Active Archive was created by ESA to provide contextual observations supporting the Venus Express mission (**?**). Near-infrared imaging can be used to sample thermal emission from the Venusian surface on the nightside (**?**). The Martian atmosphere, with its ephemeral clouds, seasonal CO2 polar ice cycles and dust storms, continues to prove popular among citizen observers, although these typically supplement the wealth of high-resolution information being returned by orbital and surface missions to the red planet. As with other planetary targets, amateur observations provide the long temporal records for the evolution of atmospheric features. Groups such as the International Society of Mars Observers (IMSO, http://www.mars.dti.ne.jp/ cmo/ISMO.html), the British Astronomical Association (BAA) and the International Mars Watch program quantitatively and qualitatively assess these amateur images.

In Figure 1 we show some examples of planetary images obtained by the amateur community.

2.1.3 Reactive to Unexpected Events

Long-term monitoring exercises prove most fruitful in response to new and unexpected events, including eruptions

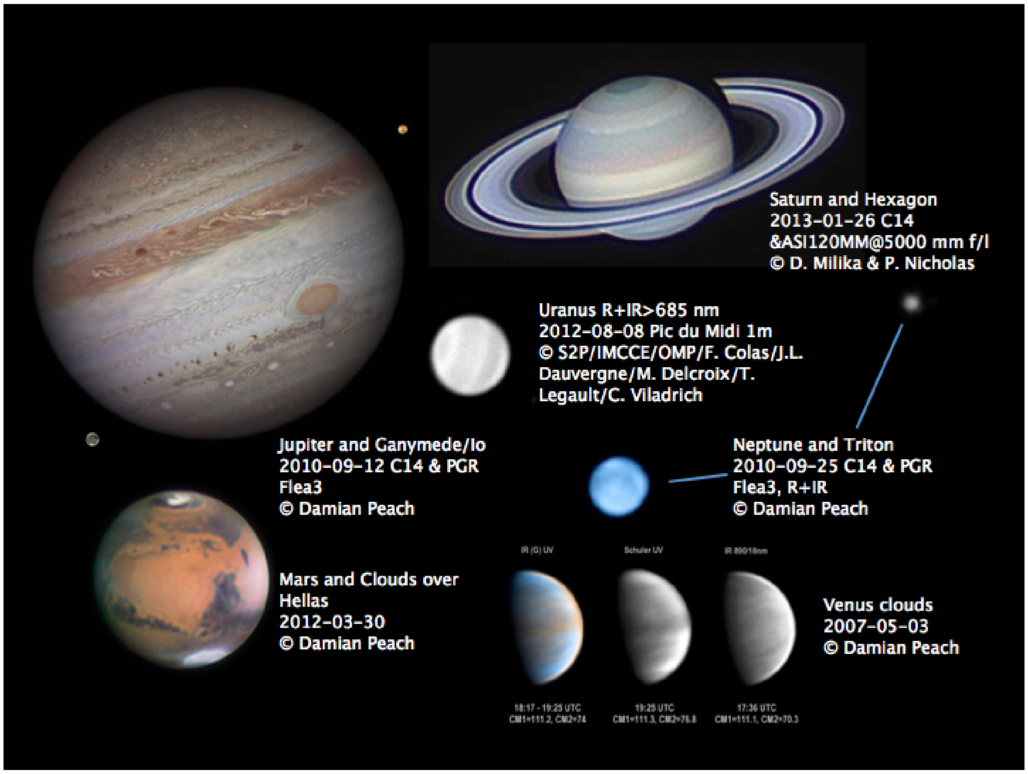

Figure 1: Examples of high fidelity images obtained by amateur planet observers.

of gigantic plumes (**??**), changes to giant ovals (**?**) and impact debris produced by asteroidal and cometary collisions (**?**). An impact scar near Jupiter?s south polar region was first discovered in imaging by Australian amateur Anthony Wesley on July 19th, 2009, and led to an international campaign of professional observations to understand the asteroidal collision that had created the scar (e.g., **??**). As observers use video imaging to capture their images (see Section **??**), those videos have also been used to identify flashes due to meteors in the jovian atmosphere. At least three flashes were confirmed between 2010 and 2012, and the light curves used to determine the sizes and frequency of objects colliding with Jupiter (e.g., **?**).

Closer to home, citizen scientists play a crucial role in the recording of rare and unpredictable events such as the fireballs from meteoroid impacts, such as the

February 2013 Chelyabinsk meteor. Video footage of the fireball and shockwave were essential to scientifically characterise the impactor and its likely origins. These reconstructed trajectories even permit the recovery of meteorites from a strewn field (i.e., when the meteor survives the intense heat of atmospheric entry and reaches the ground). These objects are the remnant debris left over from the epoch of planetary formation, and fragments left over from comets and asteroids, so their numbers, sizes and composition provide a window onto the earliest evolutionary stages of our solar system. The statistics of these impacts can only be obtained via a global network of enthusiastic citizen scientists, sharing and publicising their observations of meteors via the International Meteor Organisation (IMO, www.imo.net). Beyond Earth, transient impact flashes due to lunar impacts are recorded by video monitoring of the non-illuminated fraction of the Moon, aiming to determine the impact hazard at the lunar surface. These quantitative studies of impacts in the Earth-Moon system allow scientists to understand the meteoritic streams threading our solar system; identify previously unrecognised meteor showers; and determine the statistics of potentially hazardous encounters with this extra-terrestrial material.

Chris: We need to cover the various amateur efforts observing objects outside the solar system. Activities I am aware of are: variable star community; visual supernovae spotting; John Johnson's collaboration with an expert amateur. Can you write what you know, and suggest pointers for me to follow up please? Thanks!

The AAVS effort deserves a significant mention I think – do you know much about this? In particular we should review how they accumulate data, and ask if this is scalable/repeatable. How does it compare to

efforts in other fields, notably ecology?

2.2 Passive Observing (1 page)

While amateur astronomers have aquired a great deal of very useful data, the general population is better equipped than ever to image the sky and make that data available for scientific analysis. This has been demonstrated by two recent professionally-led studies, that made use of a largely passive observing community connected via online social networks not usually associated with astronomy.

• *The Orbit of Comet Holmes from the Photographs Uploaded to Flickr.* **?**) used N images scraped from the photo sharing website Flickr as inputs to a reconstruction of the orbit of Comet Holmes. This comet was bright enough to be visible with the naked eye in XX, 20XX, and a large number of photographs were taken of it, and uploaded to the Flickr site. **?** were able to astrometrically calibrate the images that contained enough detectable stars in the background using their automatic image registration software, `astrometry.net`. This had been enabled as a Flickr "bot," crawling over all images submitted to the `astrometry.net` group and sending the photos' owners messages showing them where on the sky their images were taken. The calibrated images trace out the trajectory of the comet over N nights, allowing a refinement of the comet's orbit of ... As the authors point out ... While in this case the photographers did not realize they were participating in a scientific study, the potential of combining powerful calibration software with large amounts of citizen-supplied imaging data is made clear.

• *Detecting Meteor Showers with Twitter.* By saving a nightly (?) log of all tweets submitted to the web service Twitter, **?**) were able to detect several new meteor showers simply by searching for the text string “meteor.” Unwitting naked-eye observers had spotted shooting stars and tweeted about them, giving rise to a detectable signal in the steam of tweets that night. The detected sample is incomplete/unlocalised/ etc... However, this work illustrates the potential both of Twitter as a communication system for connecting large numbers of observers with a science team, and of networks of unequipped observers for doing very bright object transient astronomy.

2.3 Citizen Instrumentation (Leigh, Phil to edit, 1 pages)

Instrumentation built and used by citizen scientists have increased in sophistication: complex personal observatories with large-aperture motorised telescope, multiple filters, high precision low-light cameras have become *de rigeur*. Increasingly sophisticated software is also now available to aid in the reduction and quantitative analysis of amateur imaging, as described in Section **??** below. Vibrationally-damped mounts; advanced CCD and CMOS detectors; a GOTO system to mechanically move to a set of celestial coordinates and track that location, all serve to optimise the amount of time the observer actually gets to spend looking at the sky, and paves the way for fully robotic observatories. The reader is referred to **?**) for a thorough review of instrumentation currently in use for solar system studies.

The instrumentation and software must necessarily be tuned to the phenomenon of interest, with many active observers choosing to use video monitoring to capture images in moment of excellent seeing, a technique known as lucky imaging. The best images at moments of clear seeing from the high-resolution video frames are selected, extracted and stacked together in a software programme. Some software also allows corrections of the distortions associated with telescope optics and residual atmospheric seeing. These techniques are commonly employed in observations of planetary atmospheres (e.g., Venus, Jupiter) to provide high-resolution images, and have the added benefit of providing time-sampled observations for impact flash detection (e.g., **?**). The science contributed has been driven by the instrumentation available to the citizens, but in some cases amateur observers have undertaken their own customisation work to contribute new information to the field.

We are light on "deep sky observing" in this section too. What kit is being used? Online/robotic telescopes, like the Bradford Telescope? Personal observatories, rented mountain-top space?

3 Data Processing (6 pages)

Building, instrumenting and maintaining a telescope, and then observing the night sky with it, are perhaps the most familiar activities to amateur astronomers. Meanwhile, professional astronomers spend far longer working with the data they have taken after their observing run, reducing and exploring images and spectra, and detecting and characterising objects and features. This data processing phase is an essential part of the scientific process; it results in a set of summary statistics, that can be more conveniently propagated through to the interpretation phase. That is, data processing involves distillation of data into knowledge – but stops short of the generation of understanding. What sorts of data processing have citizens been actively engaged in? We include the word "actively" here, to differentiate between the data processing that astronomers carry out, and computing jobs that can be farmed out to grids of computers owned by citizens.

Distributed "grid" computing does not fit our definition of citizen science, since it lacks the intellectual contribution element.

3.1 Image Processing (1 page)

Visual classification is by no means the only activity in which citizens have been participating. One area where visual classification and software development has translated directly into new knowledge is in the monitoring of giant planet meteorology, both using images provided by active observers, and images acquired and archived by professional facilities and spacecraft. Images are regularly qualitatively described by a number of organisations, including the British Astronomical Association?s Jupiter section[3], by a team of amateurs with substantial expertise in Jupiter's appearance (**?**). Their regular bulletins describe the changing appearance of the banded structure, the emergence of new turbulent structures and weather phenomena, and keep a record of the long-term atmospheric changes. However, recent software developments have provided a much more quantitative angle on these observations. The WinJUPOS software[4] was developed by a team of amateurs led by G. Hahn. This allows multiple images to be stacked with a correction for the rapid (once per ten hour) rotation of Jupiter or Saturn, then reprojected onto a latitude-longitude coordinate system, so that the precise positional details of atmospheric features can be determined via ?point-and-click.? By doing this over many nights surrounding Jupiter?s opposition, the team builds up enormous drift charts (tens of thousands of positional measurements) for features, ranging from the tiniest convective feature being moved by the jet streams, to the largest vortices. The positions can be extrapolated forward in time, enabling

[3] http://www.britastro.org/jupiter/

[4] http://jupos.privat.t-online.de/

targeted observations by professional observatories or even visiting spacecraft. This long-term record of Jupiter?s visible appearance by citizen scientists has proven invaluable for jovian atmospheric scientists.

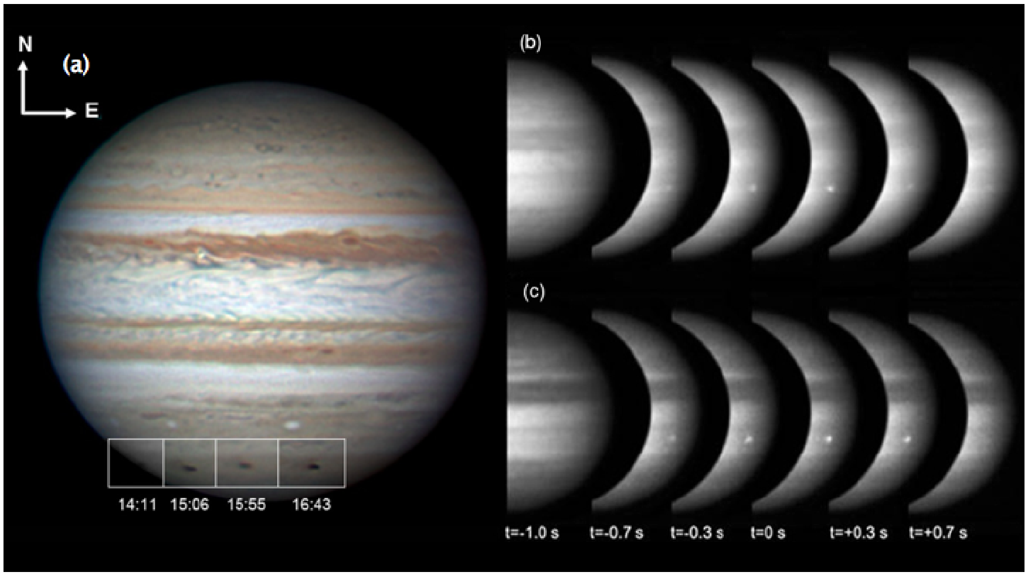

Figure 2: Citizen science contributions to monitoring of impacts in the Jupiter system. (a) Dark impact scar in Jupiter's atmosphere imaged by Anthony Wesley on July 19th 2009 (**?**). (b) The evolution of a smaller bolide impact on June 3rd 2010 at red wavelengths, also imaged by Wesley. (c) The evolution at blue wavelengths by Christopher Go, figure from **?**).

As described in Section **??**, one of the benefits of the lucky imaging technique for solar system imaging is the hours of video footage of these planets. These enormous datasets can be processed to search for impact flashes to constrain the statistics (and detectability) of collisions with a range of targets, notable Jupiter and the Moon (Figure 2). Software has been written by citizen scientists for distribution to active observers, allowing them to process their own video files to search for impacts in an automated way (e.g., Jupiter impact detections[5]

[5]http://www.pvol.ehu.es/software/

and LunarScan from the ALPO Lunar Meteoritic Impact Search[6]). This avoids the need for transfer and storage of large datasets on some centralised server. Both positive and negative detections are important to constrain the likelihood of jovian and lunar impacts.

One of the primary drivers for astrophotographers is to create artistic masterpieces, remaining true to scientific accuracy but also generating a product that they are proud to distribute to the broader community. For this reason, many active observers have ventured into the realm of software development to reduce, process and clean their astronomical images. Both commercially available software and freely-distributed home-grown codes have been used to correct for distortions and stack images, including Registax[7] and Autostakkert[8], among others. These programmes take the video files saved by the camera from the lucky imaging technique (Section **??**) and allow the user to align and stack only the clearest images. Some differences become noticeable when user preferences are introduced ? for example, use of wavelets to sharpen images can sometimes ?overprocess? the image and introduce artefacts. Colour images are put together and tuned using commercial packages like Adobe Photoshop, and the balance between red, green and blue is at the users discretion, usually with the aim of maximising contrast to show as many features as possible. When such images are used for science (e.g., point to point contrast measurements), it is often desirable to work with the raw, unprocessed images alongside the processed ones, precisely to avoid these artefacts and user biases.

It is not just active observers who have begun using sophisticated image pro-

[6]http://alpo-astronomy.org/lunarupload/lunimpacts.htm

[7]www.astronomie.be/registax

[8]www.autostakkert.com

cessing techniques. There is a growing community of citizen software specialists who devote their time to processing of raw images from interplanetary missions. Chief among these is the online forum, UnmannedSpaceflight.com, whose stated aim is to "advance public interest in, and use of, space exploration data." This includes both raw data provided by the space agencies but not published or officially released (e.g., the Cassini raw data stream of images from Saturn[9] and those from the Curiosity rover on Mars[10]), as well as those images previously released by the agencies but re-processed and colourised by dedicated citizen scientists. Depending on the quality of the original, raw images, this can sometimes produce renderings of old data in new and startlingly beautiful ways. Although not directly used for scientific enquiry, they certainly promote astronomy to the broader public.

Citizen image processing expanded enormously with the ease of high performance personal computing, and the availability of raw images from missions that the professional teams (i.e., those responsible for running the missions) simply don?t have the time or resources to process completely. From quantitative analyses of large datasets, to image processing strategies employed by active observers and armchair explorers, this field relies on software development to move forwards.

3.2 Visual Classification in Astronomy (3 pages)

As an observational science, visual classification has played a significant role in astronomical history. Despite significant advances in machine learning and computer vision, the visual inspection of data remains an important part of the

[9]http://saturn.jpl.nasa.gov/photos/raw/

[10]http://mars.jpl.nasa.gov/msl/multimedia/raw/

subject, taking advantage of the human capacity for pattern recognition. While many in the 1990s predicted that the increasing size of astronomical datasets would make such time-intensive inspection impossible, the ability of the world wide web to reach large audiences has meant the involvement of hundreds of thousands of citizen scientists in this form of data analysis. It is perhaps worth noting that this sort of project represents an inversion of the traditional model of amateur involvement in astronomer; instead of data being supplied by citizen observers for analysis by professionals data gathered by highly specialized and dedicated professional projects are passed to volunteers for analysis.

Case studies:

- Stardust@home

  While significant preliminary work had been carried out by NASA's click-workers project (see below), the project that first illustrated the potential of crowdsourcing for astronomical purposes was Stardust@home. This effort, which asked volunteers to scan through images of samples returned from Comet Wild-2 by the Stardust mission, attracted a large audience to the apparently unprepossessing task of looking for dust grains in an effort to identify samples of material from outside our Solar System. The site was built on BOSSA, an early attempt to build a generalized platform for such crowdsourcing projects (see next section), and featured a stringent test which had to be passed before classifications were counted. Despite this hurdle, more than 20,000 people took part and a variety of dust grains were removed from the aerogel for further study, at least one later proving an excellent candidate for an interstellar grain. Perhaps the most significant long-term impact of Stardust@home, though, was the demonstration that large amounts of volunteer effort were available even for relatively 'unsexy' tasks such as hunting dust grains which do not involve intrinsically beautiful images, and that, with a suitable website design and stringent testing, scientifically valuable results could be obtained.

XXXX Add references

- Galaxy morphology with Galaxy Zoo

These results directly inspired the development of Galaxy Zoo, perhaps the most prominent scientific crowdsourcing project to date. Galaxy Zoo was built on the continued importance of morphological classification of galaxies. First introduced in a systematic fashion by Hubble, and later developed by amongst others de Vaucoleurs, it remains the case that the morphology of a system is closely related to - but not entirely defined by - parameters such as colour, star formation history, dynamics, concentration and so on. In an effort to prepare for large surveys such as the Sloan Digital Sky Survey (SDSS), Lahav et al. followed especially by the work of Ball et al. developed neural networks trained on small samples of expert classifications in order to automate the process of classification,[11] arguing that the size of the then up-coming surveys left no place for manual classification.

The performance of the automatic classifiers depended on the input parameters, including colour, magnitude and size. These variables correlate well with morphology, but are not themselves morphological, and when included they dominate the classification. In particularly for galaxies which do not fit the general trends, such as spirals with dominant bulges or star-forming

[11] The Lahav papers are perhaps as interesting for their psychology as for their astrophysics, as the classifications reveal the relations between the senior classifiers employed to be experts

ellipticals, automated classifiers whether using these simple measures or more complex proxies for morphology such as texture fail to match the performance of expert classifiers. Schawinski and Nair, amongst others, spent substantial time during their time as graduate students classifying tens of thousands of galaxies.

Inspired by Stardust@home, a group led by one of the authors (Lintott) created Galaxy Zoo in 2007 to provide basic classifications of SDSS galaxies[12] Classifiers were presented with a coloured image centered on and scaled to one of more than 800,000 galaxies, and could select from one of six options: clockwise, anti clockwise and edge-on spirals, ellipticals, mergers and 'star/don't know'. Despite the initial inclusion of an easily-passed test, little knowledge was required or indeed presented to classifiers who could do something real soon after arriving on the site; this approach, in contrast to Stardust@home where a difficult test needed to be completed before authentic and useful contributions could be made, was successful in encouraging large numbers of visitors to participate. This tactic - in which both passing and sustained engagement provide substantial contribution - is illustrated in figure XXX INSERT BOX DIAGRAM XXXX which shows results from Galaxy Zoo 2. This later version of the project asked for more detailed classifications via a decision tree containing questions such as 'How prominent is the bulge?', and later iterations of the project have applied a similar approach to galaxies drawn from Hubble Space Telescope surveys including GEMS, GOODS, COSMOS and CANDELS.

[12] The original Galaxy Zoo is preserved at http://zoo1.galaxyzoo.org with the current incarnation at www.galaxyzoo.org.

XXX STATISTICS XXXX. These figures are undoubtedly impressive, but they would be meaningless if the classifications provided were not suitable for science. With sufficient effort to ensure each galaxy is classified multiple times (as many as 80 for many Galaxy Zoo images), these independent classifications need to be combined into a consensus. As discussed in later sections, this can become complex but for Galaxy Zoo a simple weighting which rewards consistency, first described in Land et al., was suffucient. Importantly, combining classifications provides not only the assignment of a label but, in the vote fraction in a particular category, an indication of the reliability of the classification. This allows more subtle biases, such as the propensity for small, faint or distant galaxies to appear as elliptical regardless of their true morphology, to be measured and accounted for (see Bamford et al.). The net result is that the Galaxy Zoo classifications are an excellent match for results from expert classification, and have produced science ranging from studies of red spirals (Masters et al.) to investigations of spiral spin (Slosar et al.).

A full review of Galaxy Zoo science is beyond the scope of this review; a recent summary is given in the Galaxy Zoo 2 data release paper by Willett et al. However, it is worth noting that many of the project's most important results have been the result not of interaction with the main interface but represent rather serendipitous discoveries made by participants. The best known such system is 'Hanny's Voorwerp' (Lintott et al.), a galaxy-scale light echo which reveals a recent (~ 100000 years) shutdown of AGN activity in IC 2497, the neighboring spiral galaxy. The discovery of the Voorwerp, first recorded in the Galaxy Zoo forum a few weeks after the project started, inspired a more systematic search for similar phenomena in other galaxies. This project, made possible by the deep engagement of Galaxy Zoo team member Bill Keel in the forum community, succeeding in finding more than forty instances of clouds which appear to have been ionized by AGN activity, in systems a third of which show signs of significant drops in activity on a timescale of tens of thousands of years.

The ability of volunteers to carry out their own research, moving far beyond the mere 'clockwork' required by the main interface, is best illustrated by the discovery of the Galaxy Zoo Peas. These small, round and, in SDSS imaging, green systems are dwarf systems with specific star formation rates (SFR per unit mass) which are unprecedented in the local Universe, matched only by high-redshift Lyman-break galaxies. Volunteers not only identified these systems, but organized a systematic search and further review of them, including using tools designed by SDSS for professional astronomers to acquire and study spectral data. The discovery of the Peas marked the first time the Galaxy Zoo team realized the potential of the community of citizen scientists the project had acquired, but it is important to note that the simpler, initial interaction provided by the main interface was necessary in order to develop that community in the first place. The participants in the citizen scientists' investigation of the Peas did not arrive on the site wanting to dig into spectra or confident of their ability to do so; these were the results of their participation. The project is, in addition to producing science, acting as an 'engine of motivation' in inspiring its participants to become more involved.

We have dwelt on Galaxy Zoo at length because it allows us to clearly see the key advantages of crowdsourcing as a solution to scientific data analysis. It is capable of scaling to the size of modern data sets to produce results of scientific value. It enables serendipity, giving individual attention to each image and allowing the identification of (and advocacy for) systems which are literally one in a million such as Hanny's Voorwerp. The large datasets it produces, and in particular the ability to quantify uncertainty, improve machine learning by providing large, rich training sets. Last but not least, crowdsourcing projects have enormous educational potential, teaching participating volunteers about science and about the process of science even when that is not the explicit goal of the project.

xxxx More references

- Surfaces of solar system bodies: Moon Zoo, Moonwatch. Saturn storms. JUPOS measurers.

  If studying galaxies remains, at least in part, a visual pursuit, then the same is certainly true of planetary science. NASA's clickworkers, which asked volunteers to identify craters on the Martain surface, lays claim to be the oldest astronomical crowdsourcing project. The consensus results matched those available from experts XXXXX CHECK XXXXX at the time, but failed to go beyond this promising start to produce results of real scientific value. More recently, interfaces inviting classifiers to look at the Moon, Mercury, Mars and Vesta have been launched and attracted significant classifications, but although preliminary results have been promising these projects have yet to produce datasets that have been used by the planetary science community in the same way that Galaxy Zoo has by the astronomical community.

- Time domain astronomy: Supernova Zoo, PlanetHunters

  If citizen science is a way of dealing with 'big data' then it is perhaps useful to remember that data sets can be big enough to be problematic in a variety of ways; the XXXX CITE XXXX describes the 'three V's' of Big Data - volume, velocity and variety. With the increasing prominence of time-domain astronomy, astronomical projects requiring immediate inspection of live data can also benefit from citizen science. While transients such as supernovae or asteroids can often be found through the use of automatic routines, visual inspection is still used by many teams as part of the process of selecting candidates for follow-up.

  The most successful attempt to use crowdsourcing to attack these problems was an offshoot of Galaxy Zoo described in Smith et al. Data from the Palomar Transient Factory XXX CITE XXX was automatically processed and images of candidate supernovae uploaded on a nightly basis; this triggered an email to volunteers who, upon responding were shown the new image, a reference image and the subtraction between the two. By analyzing the answers to a series of questions, candidates were sorted into three categories, roughly corresponding to 'probable supernova', 'likely astrophysical but non-supernova transient' and 'artifact'. The results were displayed on a webpage and used to select targets for follow-up. Despite the site attracting many fewer classifiers than the main Galaxy Zoo, it was highly effective in sorting through data with consensus typically reached on all images within 15 minutes of the initial email being sent.

  Archive projects such as the main Galaxy Zoo have the luxury of being able to be inefficient. When the appetite of volunteers for classification is high, and little of consequence is attached to finishing the task in, for example, three months instead of six, there is little incentive to direct attention efficiently. A random assignment of task to classifier thus, in most cases, suffices, but this is not true when rapid classification is important (or in the case of larger datasets). Using the supernova project's archive as a test, Simpson et al. developed a Bayesian method for assessing classifier performance; in this view, each classification provides information both about the subject of the classification and about the classifier themselves. Classifier performance given subject properties can thus be predicted and an optimum set of task assignments calculated. For systems involving tens of thousands of classifiers and perhaps a similar volume of subjects to be classified, an exact solution is computationally extremely expensive, but a MCMC approach can be used to find efficient solutions. Work by Simpson et al., as well as Horvitz et al and Waterhouse on Galaxy Zoo data, suggests that accuracy can be maintained with as few as 30% of classifications. This sort of optimization will be increasingly important for online citizen science, but a major challenge remains in incorporating this sort of analysis in live systems rather than using archive data.

- Rapid-reaction events (jovian/lunar impacts, storm/plume eruptions)
- Data mining for asteroids and TNOs.

3.3 Software development by citizens (Phil, Chris, 1 page)

JUPOS measurers: wind measurement. Impact detection.

PlanetHunters, Galaxy Zoo analysis.

Stumm at astrometry.net.

Collaborative development projects with citizens.

4 Data Modeling: Citizen Analysts (4 pages)

New understanding of the world comes from the interpretation of data in the context of a model. The modeling activity itself often has technical difficulties that computers may find hard to overcome, associated with complex and/or computationally expensive, likelihood functions. Humans, by applying their developed intuition, can often contribute a great deal to the exploration of a model's parameter space by closing in quickly on the model configurations that are fit the data well. This process can be particularly satisfying, rather like solving a puzzle. How have citizen scientists been involved in model making and data fitting in astronomy, to date?

4.1 Data Modeling in Astronomy (2 pages)

A number of web-based citizen astronomy projects include an element of data modeling as part of their set tasks. The Milky Way Project (Simpson et al. 2012) provides volunteers with a fairly flexible set of annulus-drawing tools, for annotating circularly-symmetric "bubble" features in colour-composite (24.0, 8.0 and 4.5μm) infrared images from surveys carried out by the Spitzer space telescope, which are hypothesised to have been caused by a recently-formed high mass star at the centre of the bubble. The (bubble) model in this case is simple and recognisable, making both the interface constructions and its operation relatively straightforward. The large sample of bubble models have been used to investigate the possibility of further star formation being triggered at the bubble surfaces (Kendrew et al. 2012).

Another Zooniverse project, Space Warps,[13] also involves data modeling, but not directly on the interface, which is restricted to enabling identification ("spotting") gravitational lensed images. A fraction of the community is engaged in modeling the identified lens candidates using web-based software developed and supported by the science team.[14]

Perhaps the most advanced attempt at data modeling in astronomical web-based citizen science has been the Galaxy Zoo Mergers project (Holincheck et al. 2010, Wallin et al. 2010). Here, simple N-body simulations of galaxy mergers were performed in a java applet, and the results selected according to visual similarity to images of galaxy mergers (previously identified in the Galaxy Zoo project). A key proposal in this project is that the inspectors of the simulation outputs would be able to find matches to the data more readily than a computer could, for two reasons. First is that humans are good at *vague* pattern matching: they do not get distracted by detailed pixel value comparisons but instead have an intuitive understanding of when one object is "like" another. The second is that initialising a galaxy merger simulation requires a large number of parameters to be set – and its this high dimensionality that makes the space of possible models hard to explore for a machine. Humans should be able to navigate the space using their intuition, which is partly physical and partly learned from experience gained from playing with the system. Initial tests on Arp 86 showed the crowd converging on a single location in parameter space, and that the simulated mergers at this location do indeed strongly resemble the Arp 86 system. The authors have since collected thousands of citizen-generated models for a sample of a large number SDSS merging systems (Holincheck et al, in preparation).

[13] `http://spacewarps.org`

[14] `Lens Labs`

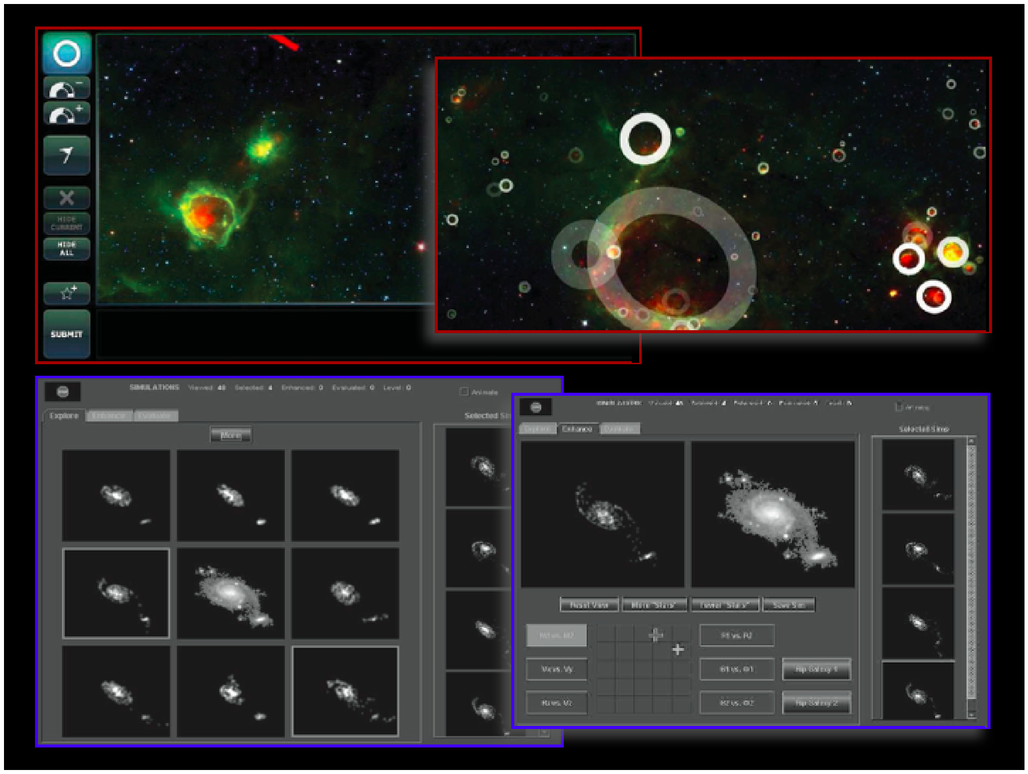

Figure 3: Examples of image modeling in web-based citizen science projects. Top row: star formation "bubble" identification and interpretation in Spitzer images in the Milky Way Project, with the annotation interface shown on the left, and some example (selected, averaged) bubbles on the right. Images from Simpson et al. (2012). Bottom row: matching N-body simulated merging galaxies to SDSS images in the Galaxy Zoo Mergers project (left), and exploring parameter space two parameters at a time to refine the models (right). Screenshots from Holincheck et al. (2010).

The above examples involved modeling infrastructure provided by either the project's developers or science teams. There have also been cases where citizens have carried out modeling analyses using their own tools, or writing their own software. For example, in the PlanetHunters project, a small group [?] of volunteers downloaded full Kepler lightcurve datasets for the best [?] community-selected candidates, and fitted transit lightcurves to them using [?]. [Chris: can you please confirm and extend this?]

Another very interesting case is that of the analysis challenges organised by the professional cosmology community. The measurement of weak gravitational lensing by large scale structure ("cosmic shear") relies on the measurement of the shapes of distant, faint galaxies with extreme accuracy. The STEP (Heymans et al. 2006, Massey et al. 2007) and GREAT (Bridle et al. 2010, Kitching et al. 2012a, 2013) blind galaxy shape estimation challenges have had an enormous impact on the field, revealing biases present in existing techniques, and providing a way for researchers outside the world of professional cosmology to participate. In particular, the GREAT08 challenge saw very successful entries (including the winner) from two (out of a total of 11) teams of researchers from outside the field (albeit still professional researchers). A companion, somewhat streamlined galaxy shape measurement challenge, "Mapping Dark Matter," which was hosted at the Kaggle website[15] (Kitching et al. 2012b). The wider reach of this platform led to over 70 teams making over 700 entries to the competition; in a comparison with the GREAT challenges, the authors found a factor of several improvement in shear accuracy over comparable previous challenges. Kitching et al. (2012b) suggest two interesting explanations for the success of the Kaggle challenge. First, the challenge was designed to be as accessible as possible, with an extensive training set of data that needed very little explanation; in this way the challenge was geared towards *idea generation*. Second, they note that the competitive nature of the challenge (a webpage leaderboard was updated in real time as entries were submitted) seemed to stimulate the analysts into improving their submissions.

[15] `http://www.kaggle.com/c/mdm`

Kaggle offers cash prizes, which will have had some effect as well (the pot was $3000 for this challenge, even if indirectly).

A second astronomical Kaggle challenge involved inferring the positions of dark matter halos based on their weak lensing effects ("Dark Worlds,"[16] Harvey et al, in prep.) This challenge attracted the attention of 357 teams, perhaps due to its larger prizes. It also sparked some debate in its forums as to the design of the challenge: the models used to generate the data, the size of the test datasets (and consequent stability of the leaderboard), the choice of leaderboard metric and so on. These issues are also of generic importance for scientists looking to crowd-source algorithm development. It is interesting to note that the Kaggle forums are a useful resource for the Kaggle development team: the citizens who are active there do influence the design of the site infrastructure and challenge rules (D. Harvey, priv. comm.).

4.2 Data Modeling in Other Fields (2 pages)

Case studies:

- Protein folding with Fold.it. Gamification as a technique.
- Other examples?

5 Citizen-led Enquiry (3 pages)

The previous sections have focused on specific, and isolated, activities in which citizens have participated. In most cases, the community's involvement has been a *contribution* to a scientific investigation, while not being involved in the design of the investigation. The most important part of any scientific investigation is the

[16] `http://www.kaggle.com/c/DarkWorlds`

question at the heart of it: what is it we are trying to find out about the world? In this section we look at some cases where the process of enquiry, the science, has been instigated or led by citizens. In principle, this is an area of great potential. Professional scientists can find it very difficult to step back from the technical details of their work, and see the bigger picture; in contrast, outsiders only see the big picture, and so we might expect them to ask some unusual, surprising and searching questions.

5.1 Individuals in action (Phil to edit, 1 page)

The constraints of funding proposals and management of research groups can often mean that professional scientists focus very narrowly on particular topics of research, using a particular technique for which they become known. Steering away from this course implies taking risks with time management, and allocation of resources to an ultimately fruitless research area can be detrimental to careers. Citizen scientists are largely free of these managerial and budgetary constraints, and are able to devote their attentions to whatever topics interest them. We propose that the creativity of citizen scientists, and the freedom to ask questions and direct scientific enquiry, is nurtured by the creation of communities. These can be online communities, with individuals contributing ideas via discussion groups and challenging one another; or they can be ready-made communities, such as school classes and family groups.

Online forums arguably provide the most direct connection between citizens and professional scientists, and have already been discussed in this review. Examples of individual successes include the ?Saturn Storm Watch? connecting Cassini?s observations of lightning emissions with active amateur observations of convective cloud structures within the giant planet atmosphere; and the tracking of the vertices of Saturn?s bizarre north polar hexagon (**?**), a 6-sided planet encircling wave that has persisted for at least 30 years but that has only recently been observed by amateur astronomers. In the first case, citizen scientists wished to identify the source of Saturn?s radio emissions. In the latter case, the long-term evolution of the hexagon vertices is being used to understand what sort of wave this is, and to identify its origins.

Astronomy as an educational tool has been used to encourage a whole new generation of citizen astronomers, and the aid of teachers in guiding and encouraging the enquiries of school children is essential for the teaching of the scientific method. The Faulkes telescope is an excellent example of citizen-led enquiry ? both student-devised and teacher-led investigations can be performed at the network of robotic observatories in Hawaii and Australia, allowing students to develop scientific questions using their own data and collaboration with other students around the world. A selection of some of these projects can be found here: http://www.faulkes-telescope.com/showcases/schools. Other observatories around the world have telescopes devoted to citizen enquiry ? the Pic-du-Midi observatory[17] in the French Pyrenees has a 0.6-m telescope devoted to amateur observers. For example, in 2013 M. Delacroix used the 1-m observatory to image details on Uranus and Neptune[18].

Case studies:

- Teacher-led science: Blackawton Bees.
- Families as research groups: Monster eyes.

[17] http://www.obs-mip.fr/pic-du-midi

[18] http://www.cloudynights.com/ubbthreads/showflat.php/Cat/0/Number/5955129/page/0/view/collaps

5.2 Facilitated research groups (Chris, Phil, 2 pages)

Case studies:

- Galaxy Zoo forum. Voorwerp, Green Peas. Lens thread: search and model.
- Planet Hunters' investigations
- Quench.
- Deep sky obs (variable nebulae etc). Amateur asteroid observations and follow-up.

6 Understanding the Citizens (2 pages)

Having surveyed some of the activities involving citizen scientists, we can now consider some questions about this community itself. Who participates in citizen science, and what motivates them?

6.1 Demographics

Who is participating in citizen astronomy? We might expect the demographics to vary with activity, and with the level of commitment required. We have some understanding of at least the former division from two studies that were carried out approximately simultaneously, one of the community participating in Galaxy Zoo, and another of the American Association of Variable Star Observers (AAVSO). **?**) surveyed the Galaxy Zoo volunteer community to investigate their motivations (Section 6.2 below), via a voluntary online questionnaire. The 11,000 self-selected Galaxy Zoo users identified as 80% male, with both genders having an approximately uniform distribution in age between their mid-twenties and late fifties. The authors point out that this is close to the US internet user age distribution, except for slight but significant excesses in numbers of post-50s males, post-retirement people of both genders, and a deficit in males under 30. The survey respondents also tended to be more highly educated than average US internet users, with most holding at least an undergraduate degree, and around a quarter having a masters or doctorate.

These findings can be compared with a survey of the members of AAVSO: Price & Paxson (2012) received over 600 responses (corresponding to about a quarter of the community of observers and society members). The education levels of the AAVSO repondents matches the Galaxy Zoo community very closely; the AAVSO age distribution is more peaked (in the mid fifties), with a similar post-60 decline but also a marked absence of younger people. The online nature of the Galaxy Zoo project seems to have increased the participation of younger (pre middle-age) people. Likewise, the Galaxy Zoo gender bias, while itself extreme, is less so than at AAVSO, where some 92% of survey respondents were male. One additional piece of information provided by the AAVSO survey is the profession of the variable star observers: most (nearly 60%) of the survey respondents were found to be working in science, computer science, engineering and education.

The Galaxy Zoo and AAVSO communities differ by more than just the nature of their activity. The smaller AAVSO community is arguably more engaged in its research, in the sense that a larger fraction of its membership is active in taking observations and contributing to analyses. It would be very interesting to know how citizen scientist motivation varied with the level of participation: dividing the Galaxy Zoo community into volunteers that contribute to the forum and those who don't could be interesting; perhaps more so would be to repeat the analysis of Raddick et al. over a wide range of projects, and look for trends there. The emergent picture thus far, however, is of a well-educated (and often scientifically trained) but male-dominated citizen science community, whose female and younger membership is likely to have been, at least in part, enabled via projects being hosted online. Continuing to lower the barriers to entry for currently under-represented demographic groups would seem both important, and within reach.

6.2 Motivation

What motivates citizen scientists? The two demographic studies referred to above also covered this question; it was the primary motivation for Raddick et al.. Having previously identified 12 categories of motivation in an earlier pilot study (Raddick et al. 2010), Raddick et al. asked the 170,000 volunteers at the time to comment on how motivated they were by each of these categories, and which was their primary motivation. The 6% who responded gave consistent answers to around 900 forum users who responded in a separate appeal, allowing us to draw conclusions about this presumably more engaged sub-population. A desire to *contribute* to science was found to be the dominant primary motivation, being selected by 40% of respondents. *Astronomy*, *science*, *vastness*, *beauty* and *discovery* were all motivation categories that were found to very important to the volunteers, while *fun*, *learning* and *community* were less important.

The AAVSO demographic survey (Price & Paxson 2012) found similar results: over a third of variable star observers cited involvement in science and research as their primary source of motivation. However, a similar number gave an interest in variable stars as theirs, perhaps reflecting a stronger focus on the science questions involved than is present in the Galaxy Zoo community. Both groups of citizen scientists are clearly quite serious in their reasons for taking part: their motivations are actually very close to those of professional scientists, as many readers of this review will recognize.

These surveys reveal a community of people many of whom may have left academic science behind as soon as they finished their education, but who still maintained a passion for astronomy and the boundaries of knowledge. Their thirst for knew information, and the desire to be part of the scientific process drives them to actively observe the night sky or to participate in analysis of large datasets.

For many people involved in citizen science, being part of a community, albeit a distributed one, that brings great enjoyment and satisfaction. With the connectivity of the internet, there is a social aspect of citizen science that unites people with shared interests. These pastimes and hobbies are often far removed from someone's "normal" life. However, *community* was not found to be a strong motivator for the Galaxy Zoo volunteers – but it is nevertheless very important for the Galaxy Zoo forum users. More recent Zooniverse projects have sought to widen participation in community discussion, hypothesizing not that it will motivate people better, but because it will help them make better contributions. Citizen scientists, like professional scientists, are primarily motivated by getting science done.

7 Ideas for the future (4 pages)

Speculation on what will be a) made possible by advances in technology and dataset size/availability, and b) feasible.

7.1 Observations and Instrumentation in the future

Today, the crucial benefits of active citizen observing are that of time and coverage, enabled by a global network of passionate and talented observers, connected and sharing their contributions via the internet. The professional community, restricted to the over-burdened resources of space observatories and ground-based facilities, do not have the capability to replace this contribution. Furthermore, would we want to? After all, the close connection between professional and amateur communities has broadened the audience for scientific discoveries, and allowed us to use astronomy to educate and enthuse the next generation of budding citizen scientists. This two-way highway is a key benefit of our present era of citizen astronomy. However, citizen scientists currently do this for fun and education, and professionals should be wary of demanding more standardised, calibrated and uniform techniques for temporal and spatial monitoring, as enthusiasm may wane if this begins to sound like 'work' rather than a hobby.

Sadly, future advances in technology may begin to widen the gap between citizen scientists and professionals once again. These may include distributed arrays of robotic 1m+ class telescopes, operating in remote regions with excellent atmospheric conditions, and trained to observe a target in a regular fashion over multiple nights (e.g., developing a consistent high-quality dataset for cloud tracking on Venus, Mars or the giant planets; or a monitoring network for meteor showers to permit 3D trajectory reconstruction). These networks would allow quasi-continuous observations over 24 hours. Certain robotic telescopes could be dedicated educational platforms, run and maintained by schools and colleges. As the images would be regularised, we could envisage automated software to track features, detect impacts, identify morphological peculiarities over time, replacing the crowd-sourced citizen analysis currently underway. But such an investment will require both international funding and considerable time and effort, and we might question whether it will ever be necessary to replace citizen scientists with automated facilities.

More exciting will be the advances in hardware available to the individual citizen observers - larger optics, more sensitive cameras, and spectral coverage extending to longer wavelengths in the infrared. Such advances in sensitivity and optics could permit citizen investigations of Uranus and Neptune; the cold and icy bodies in the distant solar system (e.g., Trans-Neptunian objects and the Kuiper Belt), and asteroids and solar system debris closer to home. Transits of extrasolar planets in front of their parent stars would be permitted from modest observatories provided they had stable conditions. New platforms might also become available to the citizen scientist, including balloon-borne observatories rising up and out of the majority of the atmospheric turbulence to provide crisper and more detailed observations of astronomical targets.

In summary, we should aim to strike a balance between increasing automation and standardisation (potentially benefiting the science investigation) and inclusivity of citizen astronomers in the scientific endeavour. After all, we may never know the source of the next breakthrough or discovery.

7.2 Classification in the future

Live data: task assignment.

Human-computer partnerships. Replacing citizens, see SN Zoo.

7.3 Data modelling in the future

Easily installed apps or browser-based tools enable outsourcing of data modelling. Operation of code, development of code. Crowd-sourcing of current detailed analyses.

7.4 Scientific enquiry in the future

Huge public databases from wide field surveys: LSST, Euclid, SKA. User interfaces designed for anyone, with social networking enabled.

Provide publishing support, see Letters.

8 The Limits of Citizen Science (3 pages)

We have argued that a critical part of 'citizen science' lies in the ability of the amateur to make an authentic contribution to science. Earlier in this part of the review, we looked forward to a richer future for such interaction, but in this section we consider the potential limits and checks on citizen science.

8.1 Limits from the data

Problems presented by data volume, and data rates. Case studies: Large samples of lenses? Transients with SKA?

Difficult/complex analyses. Microtasking only suitable for certain parts of the process?

8.2 Limits to collaboration

Collaboration between professional and citizen astronomers. Does it scale? Communication issues: forum, letters. Contrast supervisor to student, with scientist to crowd. Prospects for large collaborations? Collaborations between citizens, eventually linked to professionals?

Relationships between citizens and professionals. Mostly one-way? Examples of two-way interactions: zoo forum, solar system monitoring, spacecraft support.

8.3 Limits to access

Connections between citizen science and open data, and open publishing. Citizens reading papers: accessibility, potential barriers.

International CS. Language barriers, cultural issues.

Fast and reliable access to the internet is a pre-requisite for participation, and active observation typically requires financial investment in equipment, biasing the demographics to the developed world. An alternative is participation in educational projects such as the Faulkes telescope, providing access to world-leading equipment for educational purposes.

Does astronomy have any sort of special place in citizen science?

Breaking down of boundaries. Professionals are citizens when outside their own field. Citizens turning professional.

9 Concluding Remarks: Characteristics of Successful Citizen Science (1 page)

To emerge.

Uncompromising stance on value of citizen contributions: focus on tasks that cannot be done by machines or professionals.

Low barrier to entry, easy to make a contribution.

Emphasis on science that volunteeris to be part of.

Good communication between professionals and citizens.

High level of respect for volunteers: think of citizens as collaborators, research assistants.

The need to understand black box systems – especially if the box is full of people. But: people must be treated as ends in themselves.

Acknowledgments

PJM thanks the Royal Society for financial support in the form of a university research fellowship. CJL is grateful... LDF acknowledges... This work was supported by...

10 Literature Cited

References

1. Bridle S, Balan ST, Bethge M, Gentile M, Harmeling S, et al. 2010. *MNRAS* 405:2044–2061

2. Heymans C, Van Waerbeke L, Bacon D, Berge J, Bernstein G, et al. 2006. *MNRAS* 368:1323–1339

3. Holincheck A, Wallin J, Borne K, Lintott C, Smith A, et al. 2010. In *Galaxy Wars: Stellar Populations and Star Formation in Interacting Galaxies*, eds. B Smith, J Higdon, S Higdon, N Bastian, vol. 423 of *Astronomical Society of the Pacific Conference Series*

4. Kendrew S, Simpson R, Bressert E, Povich MS, Sherman R, et al. 2012. *ApJ* 755:71

5. Kitching TD, Balan ST, Bridle S, Cantale N, Courbin F, et al. 2012a. *MNRAS* 423:3163–3208

6. Kitching TD, Rhodes J, Heymans C, Massey R, Liu Q, et al. 2012b. *ArXiv e-prints*

7. Kitching TD, Rowe B, Gill M, Heymans C, Massey R, et al. 2013. *ApJS* 205:12

8. Massey R, Heymans C, Bergé J, Bernstein G, Bridle S, et al. 2007. *MNRAS* 376:13–38

9. Price CA, Paxson KB. 2012. *ArXiv e-prints*

10. Raddick MJ, Bracey G, Gay PL, Lintott CJ, Cardamone C, et al. 2013. *Astronomy Education Review* 12:010106

11. Raddick MJ, Bracey G, Gay PL, Lintott CJ, Murray P, et al. 2010. *Astronomy Education Review* 9:010103

12. Simpson RJ, Povich MS, Kendrew S, Lintott CJ, Bressert E, et al. 2012. *MNRAS* 424:2442–2460

13. Wallin J, Holincheck A, Borne K, Lintott C, Smith A, et al. 2010. In *Galaxy Wars: Stellar Populations and Star Formation in Interacting Galaxies*, eds. B Smith, J Higdon, S Higdon, N Bastian, vol. 423 of *Astronomical Society of the Pacific Conference Series*