International Journal of Human – Computer Studies 104 (2017) 50–63

Contents lists available at ScienceDirect

Int. J. Human–Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Task Workflow Design and its impact on performance and volunteers' subjective preference in Virtual Citizen Science

James Sprinks[a,*], Jessica Wardlaw[a], Robert Houghton[b], Steven Bamford[c], Jeremy Morley[a,d]

[a] Nottingham Geospatial Institute, University of Nottingham, Triumph Road, Nottingham NG7 2TU, UK
[b] Human Factors Research Group, University of Nottingham, University Park, Nottingham NG7 2RD, UK
[c] School of Physics & Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK
[d] Ordnance Survey, Explorer House, Adanac Drive, Southampton SO16 0AS, UK

ARTICLE INFO

Keywords:
Citizen Science
Engagement
Task workflow
Interface design

ABSTRACT

Virtual citizen science platforms allow non-scientists to take part in scientific research across a range of disciplines. What they ask of volunteers varies considerably in terms of task type, variety, user judgement required and user freedom, which has received little direct investigation. A study was performed with the *Planet Four: Craters* project to investigate the effect of task workflow design on both volunteer experience and the scientific results they produce. Participants' feedback through questionnaire responses indicated a preference for interfaces providing greater autonomy and variety, with free-text responses suggesting that autonomy was the more important. This did not translate into improved performance however, with the most autonomous interface not resulting in significantly better performance in data volume, agreement or accuracy compared to other less autonomous interfaces. The interface with the least number of task types, variety and autonomy resulted in the greatest data coverage. Agreement, both between participants and with the expert equivalent, was significantly improved when the interface most directly afforded tasks that captured the required underlying data (i.e. crater position or diameter). The implications for the designers of virtual citizen science platforms is that they have a balancing act to perform, weighing up the importance of user satisfaction, the data needs of the science case and the resources that can be committed both in terms of time and data reduction.

1. Introduction

Citizen science, also known as "public participation in scientific research" (Hand, 2010), can be described as research conducted, in whole or in part, by amateur or nonprofessional participants often through crowdsourcing techniques. Extant citizen science projects require the participant to either act as a sensor and collect data, typically 'in the wild' with an array of mobile technologies, or analyse previously collected data through internet-based Virtual Citizen Science (VCS) platforms (Reed et al., 2012). Launched in 2009, the Zooniverse (www.zooniverse.org) is home to some of the internet's most popular VCS projects, which contribute to a wide range of research, with volunteers asked to, for example, classify different types of galaxies from photographs taken by telescopes (www.galaxyzoo.org), transcribe historical ships logs and weather readings (www.oldweather.org), or mark craters found on images of planetary surfaces (www.moonzoo.org).

As a relatively new form of activity, online citizen science research has tended to be driven by concerns around the core science rather than being considered as something that can be designed to suit its user population (with some exceptions, e.g., Prestopnik and Crowston, 2012). This is perhaps ironic given the importance of the 'citizen' to the endeavour, especially as the effectiveness of a citizen science venture is related to its ability to attract and retain engaged users, both to analyse the large amount of data required, and to ensure the quality of the data collected (Prather et al., 2013). Current VCS platforms tend to require the user to carry out tasks in a very repetitious manner, the design of which are arguably driven more by the 'science case' (analogous to a 'business case' in industry) rather than any consideration of the experience of the citizen scientist (Cox et al., 2015). In the study reported here we make a first step in considering how VCS platforms can be designed to better meet the needs of the citizen scientists by exploring whether the influence of manipulating task flow predicted with similar systems would affect the rate and number of features indicated, as well as user ratings on difficulty and usability issues. We also investigate how these factors

* Corresponding author.
E-mail addresses: james.sprinks@nottingham.ac.uk (J. Sprinks), jessica.wardlaw@nottingham.ac.uk (J. Wardlaw), robert.houghton@nottingham.ac.uk (R. Houghton), steven.bamford@nottingham.ac.uk (S. Bamford), jeremy.morley@os.uk (J. Morley).

http://dx.doi.org/10.1016/j.ijhcs.2017.03.003
Received 21 March 2016; Received in revised form 13 February 2017; Accepted 7 March 2017
Available online 09 March 2017

affect the (volunteered) data's volume and accuracy by comparing it with expert judgements.

Some studies have considered motivation amongst citizen science volunteers (Reed et al., 2013; Eveleigh et al., 2014), but not considered the form of work activity itself in any depth. This may be considered remiss since forty years of research have identified a relationship between motivation, satisfaction and work design (Hackman and Oldham, 1975; Oldham and Hackman, 2010) and in recent times has been directly applied to online crowdwork (Kittur et al., 2013). Factors such as task variety, complexity and autonomy were identified as important influences on motivation and productivity, all of which can be influenced by VCS design.

We begin with a review of relevant literature on the interplay between motivation, performance and task design in the areas of Citizen Science, work design and HCI. We then introduce Planet Four: Craters – a Zooniverse citizen science project that consists of three separate interfaces that vary in task workflow design (TWD) for the marking of craters on the surface of Mars, and present a laboratory study that directly compares participants' performance and experience across the three interfaces. Finally the impact of TWD on these results, and the implications for VCS platforms and other online mechanisms, are discussed.

2. Background

2.1. Citizen Science as a distinct form of enquiry

Although VCS is a relatively new form of work, nascent research considers Citizen Science practices in their own right, beyond the scientific problems they address (Jordan et al., 2015). These studies have investigated aspects including, but not limited to: VCS typology and functionality (Prestopnik and Crowston, 2012; Reed et al., 2012); gamification (Deterding et al., 2011; Curtis, 2014; Eveleigh et al., 2013; Iacovides et al., 2013); volunteers' extrinsic motivation (Raddick et al., 2009; Reed et al., 2013; Mankowski et al., 2011; Mao et al., 2013); and volunteer behaviour (Ponciano et al., 2014; Crowston and Fagnot, 2008; Rotman et al., 2012; Nov et al., 2011). These studies, however, are predominantly concerned with the initial attraction of volunteers to a VCS platform and visceral aspects of their design, without consideration of their experience and performance in executing tasks i.e. the work that they do once they arrive, which are not easy to control. Although some recent research has considered the effect of task and judgement on volunteer performance (Hutt et al., 2013), and how they should be designed dependent on volunteer commitment (Eveleigh et al., 2014), no study to date has directly experimented with the manipulation of TWD elements to investigate their effect on volunteer behaviour, experience and scientific output. This represents an as yet missed opportunity, as TWD can be practically affected at the design stage of a project, and so it would be beneficial to understand its potential influence on the performance of citizen science.

Factors including volunteer engagement (Lahav et al., 1995), data volume (Lintott et al., 2011) and data accuracy (Hennon et al., 2014) are key to ensuring that citizen science endeavours process the large amount of data available to the standard required in order to add value to existing datasets, and as such are used to measure the success of a project. Several decades of human factors and work design research has revealed a connection between TWD factors and similar performance measures, and so the broader research question of this study is: can the lessons learnt regarding the effect of TWD on similar systems be applied to the citizen science case? If they can, whether completely or in part, it would suggest that TWD could be tailored at the design stage to improve the performance of a citizen science project. This could be achieved through an approach that practically is easier to implement compared to the considerations of existing citizen science research, regarding the extrinsic motivation provided by the science theme addressed.

2.2. Relevant insights from perceptual psychology and the design of work

VCS platforms involve processes, mechanisms and methodologies that have historically been used in other similar systems, and as such there is a wealth of research regarding their design and implementation. For example, VCS platforms, in general, ask participants to carry out a task from a discrete set of different task types (Pelli and Farell, 2010): detection (is a stimulus present/identifiable?), discrimination (the difference between two stimuli) and matching (adjusting an attribute of two stimuli until they are equal). Such tasks force the observer to make corresponding judgements (Farell and Pelli, 1999), including yes/no (is something present or not), forced choice (pick the closest match the stimuli is to a selection of pre-defined examples) and rating scales (assess the magnitude of a certain attribute of the stimuli based on a given scale). Research on these different task types in the context of image analysis shows that they affect the performance and experience of the human actor. In one of the few studies that directly considers the citizen science case, Hutt et al. (2013) compared three approaches that generate image annotations. Three forms of response were contrasted: classifications, scoring and ranking, against a ground truth estimate derived from expert annotation. Ranking was found to be the most accurate data versus expert annotation, and also the most reliable in terms of inter-participant agreement, with classification type tasks showing the lowest level of agreement. It was also found that participants produced data comparable with that of experts in terms of overall quality.

Beyond the task types and judgements required of citizen scientists, there is also the question of how the user interface presents them. Current VCS systems often require participants to do the same task(s) repetitively over a seemingly never-ending number of images, in an almost 'data entry' like manner, for no financial reward. This scenario is analogous to that found in the 1960s concerning the mechanisms of industrial work, including the fractionation and atomisation of tasks, the most well known being found on car production lines. In response to this, Hackman and Oldham (1975) developed the 'Job Diagnostic Survey' in order to better understand jobs and how they could be re-designed to improve motivation and productivity. Factors such as task variety, complexity and autonomy were identified as key to this process, all of which can be influenced in VCS design. Building on these findings, further research has found a positive correlation between motivation and task complexity (Gerhart, 1987; Chung-Yan, 2010), task autonomy (Dubinsky and Skinner, 1984; Chung-Yan, 2010) and variety (Ghani and Deshpande, 1994; Dubinsky and Skinner, 1984). Although the main body of this research concerns work over an extended period of time, which may or may not be true of volunteers regarding a citizen science platform (Eveleigh et al., 2014), the ideas act as the inspiration for this work as a form of design choice that could be applied to the VCS case.

2.3. Task Workflow Design

The concept of task workflow design is the core construct of this study. Workflow can be defined as a series of tasks that comprise an overall process, that need to be completed in order to take the work from initiation to completion. Its design can involve considerations such as the type of tasks involved, their interaction, and the sequence in which they need to be completed (i.e. sequential or parallel). These considerations can be directly related to the factors described by Hackman & Oldham (1975), and as such could influence motivation and performance. Whilst originally a concept associated with the manufacturing and business industries (Huang, 2002; Schmidt, 1998), the notion has been extended to forms of crowd sourced work due to the analogy that can be made between them. Predominantly this research has considered TWD in an overarching manner, investigating how complex processes can be deconstructed into tasks that are