Motivating Participation through Novelty Cues

COREY JACKSON, University of California, Berkeley

The voluntary nature of participation in many open collaboration platforms means platform managers and system designers need to develop strategies to attract new participants and retain existing ones. Implementing features that positively impact users' motivational states will help address challenges surrounding user recruitment and attrition. This paper reports on the design, implementation, and evaluation of one such feature – novelty cues. Prior studies have shown that novel stimuli attract and increase attention towards objects. Citizen science projects regularly ask volunteers to classify data that no other human has seen previously. An experiment was designed where volunteers were shown novelty cues. The goal was to evaluate the saliency of novelty in motivating participation among citizen science volunteers. The results showed that under most circumstances, novelty cues are effective mechanisms to increase user motivation; however, its effectiveness may be mediated by other factors such as existing system design and individual preferences for novelty.

CCS Concepts: • **Human-centered computing** → *Empirical studies in collaborative and social computing*.

Additional Key Words and Phrases: citizen science, user behavior, motivation, novelty

ACM Reference Format:
Corey Jackson. 2019. Motivating Participation through Novelty Cues. 1, 1 (December 2019), 15 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The success of many open collaboration platforms depends on voluntary contributions – in Wikipedia, people write and edit articles, in free/libre open-source software (FLOSS) projects people compile software code and documentation, and in citizen science projects people help scientists by collecting and analyzing scientific datasets. A persistent problem plaguing open collaboration is encouraging people to join, and once they join to remain active contributors. People leave open collaboration platforms for a variety of reasons, including limited social support, usability, lack of positive experiences, boredom, dissatisfaction with moderators, harassment, and time constraints [12]. In a study of Wikipedia editors, Halfaker et al. [18] found newcomers were discouraged from contributing because of the seemingly hostile nature of reverts (edits to their articles) by experienced Wikipedia editors.

Open collaboration platforms implement features hoping to increase contributions. Frequently drawing on established theories, conceptual frameworks, and methods originating in academic fields including psychology, feminist studies, cognitive science, and organizational studies, one approach to increase motivation is through positively influencing people's motivational states Kraut et al. [30]. A scan of the literature in computer supported cooperative work (CSCW) and human-computer interaction (HCI) yields a large swath of studies informed by theories and conceptual frameworks (e.g., social recognition, feedback, goal-setting) introducing both intrinsic and extrinsic rewards as

Author's address: Corey Jackson, coreybjackson@berkeley.edu, University of California, Berkeley, 102 South Hall, Berkeley, California, 94720.

XXXX-XXXX/2019/12-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

motivators to entice people to keep contributing. For instance, the Job Characteristics Model (JCM) Hackman and Oldham [16], which postulates that when jobs possess attributes such as skill variety and autonomy, task identity workers feel internally motivated, which leads to decrease absenteeism and increased work effectiveness. The JCM attributes regularly materialize in the design of tasks in open collaboration platforms with the goal of increasing individual motivation and satisfaction [1, 22, 25, 26, 28, 33, 38, 42]. Given the potential to increase motivation, open collaboration like online citizen science projects will benefit from taking stock of the varied system features and social interactions which salient in influencing people's motivational states.

Novelty is a characteristic of objects and environments frequently associated with increased human attention. Novelty theory suggests that novel characteristics of an object or environment (or stimuli) do not fit individuals' expectations or existing categorization schema, which excites a curiosity drive leading to increased attention. As people engage with new stimuli, rewards materialize, and cognitive links between rewards and the stimuli are established, causing one to continue to seek out rewards [35]. The interaction between the stimuli and rewards perpetuates a positive feedback loop giving rise to curiosity and exploratory behaviors such as increased fixation on the object [10].

This paper reports on the design, implementation, and results of an online field experiment where novelty cues were shown to users. The experiment was conducted in three citizen science projects. Citizen science projects recruit people to take part in scientific research by engaging in tasks such as choosing research questions, collecting data, conducting exploratory analysis, and analyzing datasets [11]. Citizen science projects rely on voluntary contributions making motivation a pertinent concern for most projects. One type of citizen science task asks volunteers to classify pre-existing datasets. For instance, in Galaxy Zoo, volunteers are asked to help scientists categories images of galaxies that are captured by the Hubble telescope. Galaxy Zoo volunteers answer questions about the morphological characteristics of the galaxies shown in images. Many classification projects rely on redundant classification, where multiple volunteers classify an image to provide quality responses. For volunteers, they may periodically encounter data never before seen by other humans. In line with novelty theory, it is believed that exposure to novel images (the stimuli) will excite curiosity and exploratory drives in volunteers compelling them to keep contributing. The paper seeks to address the following question: **What effect do novelty cues have on volunteer motivation?**

2 NOVELTY

Novelty is believed to induce curiosity which plays an important function in human activities such as learning [32], problem solving [19], creativity [17], and general human development [24]. Daniel Ellis Berlyne has been influential in articulating curiosity's development and influences on human behavior[5–10, 15]. Berlyne [6] disentangled curiosity and its underpinnings in humans covering forms of curiosity (i.e., epistemic and perceptual), the role of conflict in reconciliation of curiosity, and the magnitude of curiosity's arousal. Curiosity as "the condition of discomfort, due to the inadequacy of information, that motivates specific exploration" Berlyne [10, p. 2]. In articulating the foundations of novelty, Berlyne identified perceptual and epistemic curiosity. Perceptual curiosity is described as a basic form of curiosity and the driving forces for seeking out novel stimuli, predominately associated with non-humans, as it lacks goal orientation. Conversely, epistemic curiosity, causes humans to seek out and acquire information-bearing stimulation to bridge knowledge conflicts; a goal lesser observed in non-humans. Berlyne argues that conflict arises when "inadequacy of information" arises, creating tension in one's knowledge base, creating a drive to reconcile the tension. An example of this form of conflict can be found in the type of curiosity induced as a result of the introduction of new concepts to students in a classroom. As students are introduced to concepts which they cannot place squarely in existing knowledge structures, curiosity drive arises to direct efforts to reconcile the conflict between the new information and existing knowledge structures. As individuals learn, the conflict is reduced, and curiosity abates.

Berlyne's writings articulated several vital points in evaluating human responses to curiosity. First, curiosity responses are not homogeneous and depend on factors individual's preferences for curiosity and the environment in which curiosity arises. Simply some animals are more curious than others, thus, leading to different responses to stimuli. The second is that responses are governed by factors such as familiarity with the object or environment. Berlyne [6] argues that the most curiosity arousing objects are those at which there is an intermediate level of familiarity; objects for which we are entirely unfamiliar with arouses too few response tendencies (a principal operator for inducing curiosity), and objects for which we are entirely familiar with aren't interesting enough to induce curiosity. Third, repeated exposure to stimulus evoking curiosity will have diminishing effects on behavioral responses overtime, arguing these will decrease or disappear.

2.1 Novelty and its influences on human behaviors

To understand why novelty affects human behaviors, one must look towards literature unpacking its origins, neurological processes, and physiological responses. Glanzer [15, p. 303] summarizes the relationship between novelty and curiosity, stating, "When a novel stimulus affects an organism's [brain] receptors, there will occur a driving stimulus producing response called curiosity." Only recently, using technologies such as functional magnetic resonance imaging (fMRI) could neurobiologists observe and link novelty to responses in the human brain. Using fMRI imaging, Bunzeck and Düzel [13] conducted an experiment to identify which brain regions become activated when people are exposed to novel stimuli. In the experiment, participants were shown a series of stimuli (i.e., images), some of which were novel. The results of their experiment reported increased activation of the substantia nigra/ventral segmental (SN/VTA) when participants were exposed to novel stimuli. Their research also showed that responses scaled depending on the magnitude of the novelty represented in the stimuli. Other studies also implicate the hippocampus (important in the formation of memories and is also associated with learning and emotions) and the amygdala (regulating and processing emotions) as activated when humans respond to novel stimuli [13, 27, 31, 34, 40].

The SN/VTA, hippocampus, and amygdala have all been implicated in perhaps the most relevant neural mechanism for explaining the human response to novel stimuli - dopamine. Dopamine is a neurotransmitter that regulates a variety of brain functions, including movement, learning, and emotional responses to one's environment. It also mediates the brain's reward system. The primary correlate between dopamine and curiosity (and novelty) is that dopamine is the chief motivational component of reward-motivated behavior. Interacting with and anticipation of novel stimuli releases dopamine in the brain providing the sensation of reward igniting a desire to continue seeking out experiences satiating the emotion. Research attests to the surge of dopamine in stimulating excitatory processes in the brain, e.g., curiosity and resulting physiological responses – increased fixation with objects, for instance. Silvia [35] argues that as individuals engage with stimulating environments, rewards are realized (mostly intrinsic), establishing cognitive links between these rewards and engaging with the object.

Much like emotions such as anger and happiness occasions action, curiosity wontons physiological responses; these include amplified eye movements [2, 10], increased fixation with an object [37] and prolonged engagement [29, 39]. Koster et al. [29] monitored subjects' motor responses to novel and familiar images to determine the linkages between impulsivity choice and novelty. Subjects were asked to omit or perform actions by pressing a button based on the cues presented.

The authors found that motor actions were faster when cued by novel images indicating increased attention towards novelty.

Environmental factors also shape human response to novelty, adding to the complexity of measuring novelty's behavioral responses at the individual level, Berlyne [9] points out that "the most rewarding situations were those with an intermediate level of novelty, between already familiar and completely new situations." In a study in which participants were shown patterns of varying complexity, more eye movements were detected when images were novel to the participant [10], but not entirely novel. Additionally, Koster et al. [29] found response was most active when the delay between cues and action was short," meaning how novelty cues are presented may impact responses.

3 SETTING: CITIZEN SCIENCE

Citizen science involves laypersons as participants in research projects by engaging in tasks such as choosing research questions, collecting data, conducting exploratory analysis, and analyzing datasets. Citizen science projects can be in-person, online, or a combination of the two. For those that are exclusively online, most projects ask volunteers to analyze pre-existing datasets. As an example, in the Galaxy Zoo project, people are asked to view images obtained from powerful telescopes and determine whether galaxies are present and their morphologies. Once the project collects enough responses from volunteers, the classifications are handed to scientists to conduct further analysis. This paper uses classification data collected from the submission of citizen scientists in three projects on the Zooniverse platform [36].

3.1 Higgs Hunters

Astrophysicists at CERN are engaged in several scientific research projects to discover the Higgs boson particle, a project which produces millions of images that scientists struggle to analyze. To process the data, physicists, in partnership with developers from Zooniverse, launched Higgs Hunters in 2014 [3, 4]. The Higgs Hunters project shows volunteers collisions recorded by the ATLAS experiment and asks volunteers to help uncover the properties and origins of various particles by classifying detector images. The task for the volunteers is to classify the detector images for decay anomalies or appearances of off-center vertexes, which are indications that a new uncharged particle was created but then decayed into other charged particles. The classification interface is shown in Figure 1 (top,left). To record a classification, volunteers click on "Off-center vertex" on the right-hand side of the window, then mark the location of the vertex and how many tracks appeared.

3.2 Asteroid Zoo

Asteroid Zoo supports the National Aeronautic Space Administration's (NASA) Near-Earth Object Observation Program (NEOO), whose goal is to discover and catalog near-Earth objects (NEOs). The Catalina Sky Survey (CSS) is a program using powerful telescopes to produce mappings of the sky and produces millions of images per year.

The Asteroid Zoo project was launched to help scientists filter NEOs. Volunteers are asked to help scientists search for NEOs by reviewing some of the images. Listed on the site are several scientific goals guiding the efforts of the scientists – identify near-earth asteroids, catalog objects for asteroid mining, study the solar system, and collect training images for future machine classification. The classification task requires volunteers to examine a temporally linked set of images from the sky survey to identify moving objects, which could be indicative of NEOs such as asteroids or potentially hazard space artifacts. The classification interface is shown in Figure 1 (top, right). While classifying volunteers have access to a digital toolbox with features to help volunteers inspect the data object,

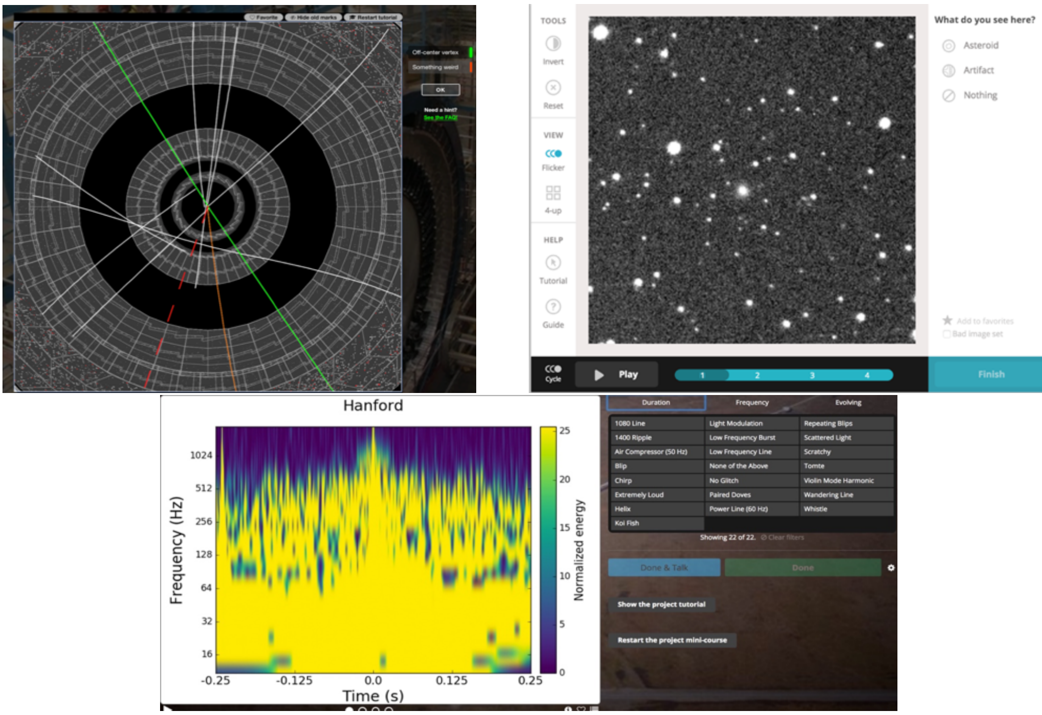

Fig. 1. The classification interface of Higgs Hunters (top, left), Asteroid Zoo (top, right), and Gravity Spy (bottom).

including inverting the image, changing the resolution for viewing, and cycling through several data objects in the series.

3.3 Gravity Spy

Launched in October 2016, the Gravity Spy project asks for the help of volunteers in categorizing datasets generated from the Laser Interferometer Gravitational-Wave Observatory (LIGO) collaboration. To detect gravitational waves, LIGO uses interferometers, which detect interference in light patterns. The machines are susceptible to noise (or glitches) surrounding the instrumentation. The instrument for recording gravitational waves produces a spectrogram. Occasionally, events such as birds chirping can be recorded by the interferometers and represented in the spectrogram images. To detect gravitational waves, scientists need to identify and remove the noise from datasets. In the Gravity Spy project, volunteers determine whether the glitches represented in the spectrogram share a similar morphological structure to a set of pre-defined categories. Depicted in Figure 1 (bottom) is a glitch and the classification interface. Scientists use the labeled dataset to isolate noise in potential gravitational wave signals.

4 METHODS

Described in this section are the experiment, data, and analysis.

4.1 Experiment Procedure

Each project described above periodically uploads new subjects to the Zooniverse platform. Thus, during some periods, there are new subjects that have been classified. A novelty cue was designed to test the effects of novelty on volunteer motivation, informing participants when they were the first volunteer to see the subject. The novelty cue reads, "Guess what! You're the first to see this subject."

The Zooniverse queuing algorithm, which is uniform across projects on the platform, serves a packet of images on the back-end when volunteers first navigate to the classification interface. Volunteers then cycle through images in their queue (upon classifying the images more are added to the queue). Volunteers classify one image at a time. When an image is loaded and has no classifications on the Zooniverse system, the novelty cue is displayed contiguously while the volunteer classifies the image. Once the classification has been submitted, a new classification is rendered. Volunteers do not know how many new images are in their queue and can stop contributing at their discretion and without consequence.

At the time the experiments in Higgs Hunters and Asteroid Zoo were run, Zooniverse did not support customization of user control variables such as assignment to treatment or control group so the intervention was administered to all volunteers who had new data objects in their queue and the experiment ran until all new data had been exhausted. Since the randomization of volunteers into treatment and control was not achievable (a solution is discussed below), the Higgs Hunters and Asteroid Zoo experiments were quasi-experiment. The Gravity Spy experiment was conducted after the first two experiments were completed, and a new system for customizing experiences was deployed, making random assignment was achievable.

4.2 Data Collection

Database dumps from each project were acquired with all the classification data. The system recorded each volunteers' classification response, whether the novelty cue was displayed, and a timestamp indicating the exact date, and time the response posted to the system.

Since randomly assigning volunteers to treatment and control groups was not possible in Higgs Hunters and Asteroid Zoo and controlling for the diminishing volume of new subjects added to the project was necessary, careful, and equitable selection of the treatment and control was required. Because Higgs Hunters and Asteroid Zoo are quasi-experiments, sessions in treatment and control *periods* (as opposed to volunteers assigned to a treatment or control group) were identified and analyzed. Table 1 shows the experiment period and the selected treatment and control periods. All sessions were plotted to visualize sessions containing the treatment. The treatment was selected by identifying a period that excluded holidays and contained approximately the same number of volunteers as a control. Both control and treatment periods were multiples of 7 days to include all days of the week equally. The selected treatment period January 21–February 3 for Higgs Hunters and June 23–July 07 was chosen for Asteroid Zoo to represent the treatment. Only sessions from volunteers who had their first session during the experiment period were used to eliminate possible influences from prior experiences on the system; however, not all sessions during this period had a novelty cue (an issue discussed below). The selected control period was January 21–February 3 for Higgs Hunters and June 23–July 07 was chosen for Asteroid Zoo. Since Gravity Spy was a randomized experiment, there was no need to select treatment and control periods.

4.3 Data Analysis

A common assumption in the analysis of experiments is that all participants in the treatment were treated. The random administration of the treatment in these experiments means that not all

Project	Experiment Period (Days)	Treatment Period	Control Period
Higgs Hunters +	Jan. 1- Feb. 11 2015 (41)	Jan. 21–Feb. 3	Jan. 8–14 & Feb.2–11
Asteroid Zoo +	Jun 23 -July 21 2014 (38)	Jun. 23–Jul. 07	Jul. 08–Jul. 21
Gravity Spy ++	May 05- May 21 2017 (16)		

Table 1. The treatment and control periods for each project. + indicates a quasi-experiment with non-randomized assignment to a treatment/control. ++ indicates randomized assignment to treatment/control.

individuals who were assigned to the treatment group were treated - some volunteers may end their session before seeing a novelty message. Randomized controlled trials in clinical research also suffer from similar challenges as patients may subsequently withdrawal from the treatment, deviate from the protocol, or become non-compliant. In such cases, an approach called the intention to treat (ITT) is used to analyze the data [21]. The logic of the intention to treat is to create treatment and control groups that are hypothetical identical by taking computing the statistics for those non-compliant participants into a separate control group in the treatment and control. Depicted in Figure 2 is the intention to treat (ITT). The ITT was necessary to analyze the data since not all of the sessions selected for the treatment group actually received the treatment. After the hypothetical groups were constructed, t-tests were used to determine whether significant group differences existed between the treatment and control groups. The experiment data were analyzed using R Studio.

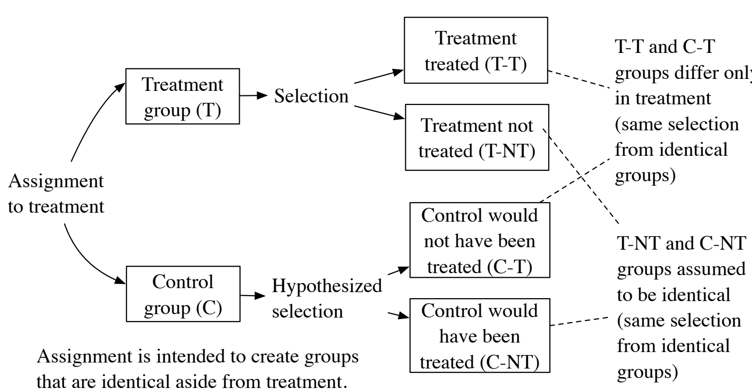

Fig. 2. Logic of intention to treat analysis.

5 RESULTS

Shown in Table 2 are the descriptive statistics for each project during the observation periods. These statistics include data that were not selected for the treatment and control periods. These data are helpful to understand the activity in the project during the time of the launch of the experiments in each project. There are noticeable differences in some of the statistics reported (See Table 2). First, while having approximately congruent observation periods there were fewer volunteers and sessions represented in Higgs Hunters (volunteers = 1,867, sessions = 3,096) when compared to Asteroid Zoo (volunteers = 3,741, sessions = 8,496). The number of volunteers in Gravity Spy (volunteers = 319, sessions = 1,097) is even smaller if controlling for the number of days in the observation period. This discrepancy likely comes from the time frame in which the experiments launched. The observation period in Higgs Hunters began on January 1st 2015 and lasted until February 11th 2015 (44 days after the project officially launched). The observation period for Asteroid Zoo began June 23rd, 2014 and ended July 21st 2014 (beginning the day the project officially launched) and May 5th to May 21st 2017 in Gravity Spy (6 months after the project launched).

Project	Volunteers	Sessions	Classifications	Novelty Cues
Higgs Hunters +	1,867	3,096	195,970	28,577
Asteroid Zoo +	3,741	8,496	392,750	45,195
Gravity Spy ++	319	1,097	82,895	4,622

Table 2. Descriptive statistics for each project.

The selection of the treatment and control periods were for investigation are shown in Table 2. The selection of the treatment and control periods were so that they are temporally adjacent and had a similar number of volunteers contributing. The data reveal that during treatment periods in Higgs Hunters and Asteroid Zoo, volunteers executed more classifications than those in the control; 19.2 more in Higgs Hunters and 5.3 more in Asteroid Zoo. However, in the treated Gravity Spy sessions resulted in 4.2 fewer classifications executed.

Since the system could not change to ensure every session during the treatment actually saw new data. During some sessions in the treatment period, no novelty cues were displayed to volunteers (treatment – not treated). A possible confound, however, is that seeing novelty cues becomes a matter of chance since some and not all data are new. This means that for treatment periods, rather than the cue causing sessions to be longer, a longer session could increase the chances of seeing a novelty message.

To mitigate the confound described in the previous section, the same intention to treat analysis conducted on this dataset. The ITT analysis used the population statistics from the treatment group to infer statistics, i.e., mean and standard deviation for a hypothetically treated control group (PC-T). A comparison of the hypothetically treated control group against the session in the treatment that was actually treated (T-T) occurred.

Reported in Table 4 are the results of the intention to treat analysis. When teased out from the entire experiment period, the T-T group had noticeable increases in the mean number of classifications executed during a session. It is worth noting that in applying intention to treat, the hypothesized control mean values are higher than those observed by simply taking the mean of the

Project	Type	Date (days)	Volunteers	Sessions	μ Class. (σ)
Higgs	PC	Jan. 8–14 (7) & Feb.2–11 (7)	183	285	27.3 (43)
Hunters	T	Jan. 21–Feb. 3 (14)	217	356	46.4 (74.7)
Asteroid	PC	Jul. 08–Jul. 21 (14)	1,367	6,136	23.20 (33.64)
Zoo	T	Jun. 23–Jul. 07 (14)	3,905	13,279	28.60 (51.13)
Gravity	C	May 05–21 (14)	229	909	55.29 (95.42)
Spy	T		193	824	51.04 (77.51)

Table 3. Descriptive statistics for the three projects during the selected experiment periods. Higgs Hunters and Asteroid Zoo were designed and analyzed as quasi-experiments where control groups were derived from periods when no novelty cues were shown. The second column Type indicates whether the observation was a Pooled Control (PC), Control (C), or Treatment (T) period.

	Higgs Hunters		Asteroid Zoo		Gravity Spy	
	Sessions	**μ Class. (σ)**	**Sessions**	**μ Class. (σ)**	**Sessions**	**μ Class. (σ)**
Treatment (T)						
Treated (T-T)	223	64.7 (84.6)	7,000	41.28 (65.1)	484	70.52 (84.87)
Not treated (T-NT)	133	15.7 (37.9)	6,279	14.46 (20.64)	308	23.17 (54.71)
Control (PC)						
Hypothetical treated (PC-T)	179	34.4 (44.4)	3,235	30.04 (40.44)	537	77.51 (144.6)
Hypothetical Not treated (PC-NT)	106	15.7 (37.9)	2,901	14.46 (20.64)	434	23.17 (54.71)

Table 4. Descriptive statistics for treatment and hypothetically derived control groups. The statistics comparison is made vertically for each project on the mean number of classifications, that is the T-T and PC -T dimensions.

control period; evidence that the analysis doesn't simply attempt to suppress values in the control period. The differences seen by comparing the treatment treated (T-T) and the hypothetically treated (PC-T) groups were only statistically significant in Higgs Hunters and Asteroid Zoo. Volunteers executed 34.4 (σ = 44.4) more tasks in Higgs Hunters at $t(400)$ = 4.35, p <.001 and 10.24 (σ = 44.4) more in Asteroid Zoo at $t(10232)$ = 8.25, p <.001. However, volunteers in T-T group in Gravity Spy contributed 7.22 fewer classifications when exposed to the novelty treatment. However, the difference was deemed to be non-significant at $t(1018)$ = -1.21, p = 0.19.

5.0.1 Novelty and First Session Behaviors. For many newcomers, their first and last session occurs in the same sitting. Thus, strategies to encourage contributions during this period are valuable. To determine whether reinforcing novelty cues could apply equally to newcomers, an analysis of classifications executed by this subset of volunteer contribution histories occurred. Again, since not all volunteers who contributed during the treatment period received the intention to treat analysis application.

The results of the analysis are in Table 5 and show newcomers executed significantly fewer classifications in the PC-T in Higgs Hunters and Asteroid Zoo than in the T-T group. The average number of tasks executed in T-T for Higgs Hunters was 51.9 (σ = 69.3) compared to 24.9 (σ = 34.4) in the PC-T. The difference of 27 classifications significant at t(247) = 3.8, p <.001. In Asteroid Zoo, newcomers in the PC-T executed, on average, 27.67 (σ = 36.75) classifications while newcomers in the T-T executed only 3.75 (σ = 9.7) more at 30.23 (σ = 55.1). The difference observed between newcomers in the PC-T and T-T was non-significant $t(3,331)$ = 0.82, p = 0.28 indicating the novelty cues were ineffective in encouraging newcomers to execute more classifications. The results for newcomers in Gravity Spy, revealed similar outcomes, that novelty cues have no effect on the number of tasks volunteers executed. In the T-T group volunteers executed fewer classifications (μ = 108.42, σ = 118.56) compared to the PC-T group (μ = 114.62, σ = 103.44) the 6.21 difference was not significant at $t(124)$ = - 0.23, p = 0.39.

5.1 Preventing Dropouts

The literature on participation patterns reveals that most volunteers in citizen science *dabble* and contribute in a single session. One potential benefit of exposing volunteers to novelty cues is that they will be excited about the possibility of seeing more novel data in the future and will be motivated to return to the project to continue classifying. One way to measure excitement to return to the project is to determine whether novelty cues decrease the dropout rate for volunteers

	Higgs Hunters	Asteroid Zoo	Gravity Spy
		μ **Class. (σ)**	
	Treatment (T)		
Treated (T-T)	51.9 (69.3)	30.2 (55.1)	108.4 (118.6)
Not treated (T-NT)	7.1 (10.9)	7.7 (15.3)	13.9 (21.4)
	Control (PC)		
Hypothetical treated (PC-T)	24.9 (34.4)	27.67 (36.75)	114.6 (103.4)
Hypothetical Not treated (PC-NT)	7.1 (10.9)	7.7 (15.3)	13.9 (21.4)

Table 5. Newcomer classifications using the intention to treat analysis.

exposed to the novelty cues. However, since novelty cues do appear to encourage newcomers to contribute more under some conditions, novelty cues might also play a role in decreasing the number of volunteers drop-outs. This section focuses on volunteer behaviors after first sessions (S) in which a volunteer was treated (T) or non-treated (N-T). The proportions of volunteers who were treated in their first session S-T and not treated S-NT in Table 6.

	S-T	S-NT	χ^2 (p-val.)
	Higgs Hunters (N= 463)		
Drop-out	96 (71%)	266 (81%)	5.9 **
Sustained	40 (29%)	61 (19%)	
	Asteroid Zoo (N=4,369)		
Drop-out	1,390 (45%)	752 (57%)	50.6 ***
Sustained	1,666 (55%)	561 (43%)	
	Gravity Spy (N=147)		
Drop-out	42 (66%)	60 (73%)	0.47
Sustained	22 (34%)	23 (27%)	

Table 6. A contingency table showing the proportions of volunteers who are treated and return after their first session (S) being treated (T) or not treated (N-T). A chi-squared test of proportions was conducted to determine equality of proportions. Significance levels *** p <0.001, ** p<0.01, * p >0.05.

For each project, the percentage of volunteers returning for an additional session increased when seeing a novelty cue during their first session increased compared to the population of volunteers not receiving a novelty cue during the first session. In all projects, the proportion of returning volunteers was greater than in the S-T than the proportion in S-NT. A χ^2 test of proportions was performed to determine whether the proportions were equal in both groups. In Higgs Hunters, the results of the tests revealed significant differences $\chi^2(1) = 5.9$, $p < 0.01$ as did the test for Asteroid Zoo; $\chi^2(1) = 50.6$, $p < 0.001$. Thus, one can conclude for Higgs Hunters and Asteroid Zoo the proportion of volunteers returning in the S-T was significantly different from those in the S-NT. In Gravity Spy, although in sessions where volunteers were exposed novelty cues performed slightly better than volunteers not receiving the novelty cues, the difference was not significant, $\chi^2(1) = 0.47$, $p = 0.49$.

5.2 Returning Sooner with Novelty Cues

If volunteers' motivational requirements are satisfied, they may be excited to return to the project sooner than other volunteers. A comparison between sessions was conducted to determine whether exposure to novelty during the first session precipitated a faster return to the project for those volunteers who did eventually return for a second session. Calculating the time to return was simply by subtracting the time after a volunteer's first session concluded and the start of their second session. The results are in Table 7 and show the time (in days) between a volunteer's first and second session. In all projects, volunteers who received treatment during their first session (S-T) took fewer days to return than those who did not receive treatment during their first sessions (S-NT). The results of an independent samples t-test revealed significant difference in the days to return for S-T (μ = 4.11 σ =15.38) and S-NT (μ = 5.5, σ = 19.6) in Higgs Hunters $t(269.15) = 1.98$, $p < 0.05$; and Asteroid Zoo (S-T μ = 5.53, σ = 19.38) (S-NT μ = 9.7, σ = 29) at $t(863.77) = 3.4$, $p < 0.001$. However, the return time in Gravity Spy was not significant $t(155.35) = -0.22$, $p = 0.82$.

		Higgs Hunters	**Asteroid Zoo**	**Gravity Spy**
Time (in days) between the	S-T	7.1 (17.4)	5.5 (19.6)	6.8 (10.8)
first and second sessions	S-NT	12.1 (25)	9.7 (29)	6.5 (15.1)
Independent samples t-test		4.56*	3.4***	-0.22

Table 7. The mean time (in days) to beginning the second session for volunteers who were treated T or not treated (N-T) during their first session (S). Significance levels *** p <0.001, ** p<0.01, * p >0.05.

6 DISCUSSION

The analysis conducted in this paper show that novelty cues are a useful mechanism to motivate volunteers to contribute to citizen science projects. Earlier studies support these findings conducted in laboratory settings, which showed that when exposed to novel stimuli, people tend to explore more than people exposed to non-novel stimuli [37]. A summary of the experiment results is included in Table 7 and show positive effects for sessions in which volunteers were treated with novelty cues in Higgs Hunters and Asteroid Zoo. In these projects' volunteers executed 34.4 (+88%) and 10.24 (+37%) more classifications than sessions included in the control groups. The results in Gravity Spy, however, showed that for the total population, sessions with the administration of a novelty cue resulted in a decrease (-7.22) in classifications executed. The difference, however, was not significant.

This analysis showed that newcomers (the population most likely to drop out) could be motivated to contribute more prior to their departure. In volunteers' first sessions, our experimental manipulation increased the number of classifications from 24.9 to 51.9, again doubling the number of classifications contributed.

In addition to motivating classification in sessions, exposure to novelty also improves retention. Given the voluntary nature of citizen science projects, most projects struggle to retain contributors. Novelty cues had a positive effect on retention in Higgs Hunters and Asteroid Zoo. Novelty cues decreased the percentage of dropouts observed 81% vs. 71% in Higgs Hunters and 57% vs. 45% in Asteroid Zoo; while there was an improvement in Gravity Spy 73% vs. 66%, the difference in proportion was not significant. Additionally, volunteers returned to Higgs Hunters and Asteroid Zoo when shown a novelty message in a prior session – 5 days sooner in Higgs Hunters and 4.2 days sooner in Asteroid Zoo.

While novelty cues did not appear to influence the behaviors of Gravity Spy participants significantly, it is suspect that elements of gamification not found in Asteroid Zoo and Higgs Hunters may take precedence as the primary motivational driver. Participation in Gravity Spy is scaffolded with volunteers being able to access new levels as they classify subjects and are accurate. This discrepancy gives rise to questions of how many motivators can be operational at one time.

	Total population	Newcomers	Retention
Higgs Hunters	***Supported***, volunteers made 34.4 more classifications during treated sessions	***Supported***, volunteers made 27 more classifications during treated sessions	***Supported***, there were fewer dropout and faster returns
Asteroid Zoo	***Supported***, volunteers made 10.2 more classifications during treated sessions	***Not Supported***, non-significant difference in contribution	***Supported***, there were fewer dropout and faster returns
Gravity Spy	***Not Supported***, non-significant difference in contribution	***Not Supported***, non-significant difference in contribution	***Not Supported***, non-significant difference in contribution

Table 8. A summary of the experiment outcomes.

6.1 Novelty and motivation

Novelty is an intrinsic motivator causing exploratory drive, i.e., curiosity giving rise to cognitive rewards systems facilitated by dopamine activity in the brain implicating humans in a feedback loop invoking exploratory drive behaviors. While this research contributes to the literature on motivation in open collaboration platforms, and citizen science projects specifically, it also establishes linkages between people's experiences online and their behavioral responses more generally. Empirical work on novelty has appeared across a diverse set of contexts from consumer behavior [20], information search and retrieval [41], game design [14, 39], and many several others [2, 23, 29]. The results tend to suggest that novelty is associated with behavioral responses such as increased attention to stimuli containing novel characteristics.

In the context of these results, the curiosity derived from engaging with novel stimuli appear to not only be dictated by the ability of designers to alert people to novelty, but also a host of individual, cultural, and environmental attributes. A person's individual's baseline preferences for novelty or whether they possess a "novelty trait." should dictate responses. Observed effects on human behaviors are not unitary as some individuals are more impacted by novelty. Those having no "novelty trait" are unlikely to realize the cognitive rewards associated with exposure to novelty stimuli. Previous studies have pointed to low percentages of the population who possess the trait and are thus expected to be behaviorally influenced by novelty. Gallagher [14] suggests that neophiliacs (the term used to describe the population of people excited by novelty) represents approximately 15 percent of the U.S. population.

Finally, novelty's effect on human motivation is not indissoluble. Early postulations establishing links between novelty and curiosity drive suggests that while human responses to novelty may fluctuate, there is a strong tendency for a reduction of curiosity drive. This reduction would, in turn, lead to decreases in cognitive and physiological responses, overtime [5, 35]. For instance, Murty et al. [31] observed decrementing responses as that as participants were exposed to novel, but unfamiliar stimuli suggesting additional work is needed to uncover how people habituate to novelty cues and motivators more generally.

7 CONCLUSION

Novel attributes associated with objects and environments give rise to a curiosity drive, which facilitates the maintenance of interaction (or motivation) to continue exploring. This research showed that implementing novelty cues may be a viable system feature to increase positively influence volunteers' motivational states. While most hypotheses across projects were supported, one finds additional complexities in responses to novelty may depend on many factors, including individual personality differences, the strength of novel attributes, and periodicity of novelty exposure.

7.1 Limitations

There are four limitations. First, is the experimental design itself: a quasi-experiment. Given the wholesale introduction of the treatment to the population, this design was the only feasible way to analyze the data from the system. A true randomized controlled experiment would have been preferred, but the system was not able to support a randomized assignment. Second, there is the possibility that the effects of the treatment last longer than the single session that was analyzed. For example, the effects of a message seen during a volunteer's first session might prompt the volunteer to return for a second session or have continuing impacts in that session.

Third, there was no examination of whether seeing multiple novelty cues in a session increases the impact of the treatment or if they instead become annoying and a distraction. Future work can explore the temporal dimensions of the effect. Finally, while an experiment provides strong evidence for a causal relationship between a treatment and an outcome, there is a trade-off for the richness of data. For example, our analysis does not include possible differences in volunteers that are not captured by the system (e.g., demographics or education). Nor does it provide the kind of rich data needed to illuminate the mechanism of the effect, that is, why volunteers found novelty to be motivating. Prior research found that volunteers report being so motivated, but understanding in more detail exactly why is a question for further in-depth investigation. More research is needed to conclude, specifically, what drives the observed phenomenon: novelty seeking, discovery, curiosity, sensation seeking, neophilia, some other term, or maybe some combination of existing drivers.

ACKNOWLEDGMENTS

TBD

REFERENCES

[1] Obinna Anya. 2015. Bridge the Gap! What Can Work Design in Crowdwork Learn from Work Design Theories?. In *CSCW '15*. ACM Press, New York, New York, USA, 612–627.
[2] Adrien Baranes, Pierre-Yves Oudeyer, and Jacqueline Gottlieb. 2015. Eye movements reveal epistemic curiosity in human observers. *Vision Research* 117, C (Dec. 2015), 81–90.
[3] A J Barr, A Haas, and C W Kalderon. 2018. Citizen scientist community engagement with the HiggsHunters project at the Large Hadron Collider. *Research for All* 2, 2 (July 2018), 359–373.
[4] A J Barr, A Haas, and C W Kalderon. 2018. 'That looks weird' — evaluating citizen scientists' ability to detect unusual features in ATLAS images of LHC collisions . *arXiv* (Aug. 2018), 1–17.
[5] D E Berlyne. 1950. NOVELTY AND CURIOSITY AS DETERMINANTS OF EXPLORATORY BEHAVIOUR. *British Journal of Psychology* 41, 1-2 (Sept. 1950), 68–80.
[6] D E Berlyne. 1954. A THEORY OF HUMAN CURIOSITY. *British Journal of Psychology* 45, 3 (Aug. 1954), 180–191.
[7] D E Berlyne. 1957. Conflict and information-theory variables as determinants of human perceptual curiosity. *Psychol Rev.* 53, 6 (1957), 399–404.
[8] D E Berlyne. 1957. Uncertainty and conflict: a point of contact between information-theory and behavior-theory concepts . *Psychol Rev.* 64, 6 (Dec. 1957), 329–339.
[9] D E Berlyne. 1960. Conflict, arousal, and curiosity. (1960).
[10] D E Berlyne. 1966. Curiosity and Exploration. *Science, Technology, & Human Values* 153, 3731 (July 1966), 25–33.
[11] Rick Bonney, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk, and Candie C Wilderman. 2009. *Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. A CAISE Inquiry Group Report.* Technical Report.
[12] Peter Bae Brandtzæg and Jan Heim. 2008. User loyalty and online communities: why members of online communities are not faithful. *CHI'07* (2008).
[13] Nico Bunzeck and Emrah Düzel. 2006. Absolute Coding of Stimulus Novelty in the Human Substantia Nigra/VTA. *Neuron* 51, 3 (Aug. 2006), 369–379.
[14] W Gallagher. 2011. New: Understanding our need for novelty and change.

[15] Murray Glanzer. 1958. Curiosity, exploratory drive, and stimulus satiation. *Psychological Bulletin* 55, 5 (Sept. 1958), 302–315.

[16] J Richard Hackman and Greg R Oldham. 1976. Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance* 16, 2 (1976), 250–279.

[17] Lydia Paine Hagtvedt, Karyn Dossinger, Spencer H Harrison, and Li Huang. 2019. Curiosity made the cat more creative: Specific curiosity as a driver of creativity. *Organizational Behavior and Human Decision Processes* 150 (Jan. 2019), 1–13.

[18] Aaron Halfaker, Aniket Kittur, and John Riedl. 2011. Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. In *Proceedings of the 7th international symposium on wikis and open collaboration*. ACM, 163–172.

[19] Jay H Hardy III, Alisha M Ness, and Jensen Mecca. 2017. Outside the box: Epistemic curiosity as a predictor of creative problem solving and creative performance. *PAID* 104, C (Jan. 2017), 230–237.

[20] Elizabeth C Hirschman. 1980. Innovativeness, Novelty Seeking, and Consumer Creativity. *Journal of Consumer Research* 7, 3 (1980), 283–295.

[21] Sally Hollis and Fiona Campbell. 1999. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 319, 7211 (1999), 670–674.

[22] Robert Houghton, James Sprinks, Jessica Wardlaw, Steven Bamford, and Stuart Marsh. 2019. A sociotechnical system approach to virtual citizen science: an application of BS ISO 27500:2016. *WebSci 2013 Workshop: Creativity and Attention in the Age of the Web.* 18, 01 (Jan. 2019), 1–25.

[23] Lucas M Jeno, Vigdis Vandvik, Sigrunn Eliassen, and John-Arvid Grytnes. 2019. Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. *Computers & Education* 128 (Jan. 2019), 398–413.

[24] Jamie Jirout and David Klahr. 2012. Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review* 32, 2 (June 2012), 125–160.

[25] Nicolas Kaufmann, Thimo Schulze, and Daniel Veit. 2011. More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk.. In *Proceedings of the 21st European Conference on Information Systems.*

[26] Aniket Kittur, Jeffrey V Nickerson, Michael S Bernstein, Elizabeth Gerber, Aaron D Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *CSCW '13: Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 1301–1318.

[27] Robert T Knight. 1996. Contribution of human hippocampal region to novelty detection. *Nature* 383, 6597 (Sept. 1996), 256–259.

[28] Masatomo Kobayashi, Shoma Arita, Toshinari Itoko, Shin Saito, and Hironobu Takagi. 2015. Motivating Multi-Generational Crowd Workers in Social-Purpose Work. In *CSCW '15*. ACM Press, New York, New York, USA, 1813–1824.

[29] Raphael Koster, Tricia X Seow, Raymond J Dolan, and Emrah Düzel. 2016. Stimulus Novelty Energizes Actions in the Absence of Explicit Reward. *PLoS ONE* 11, 7 (July 2016), e0159120–10.

[30] Robert E Kraut, Paul Resnick, and Sara Kiesler. 2011. *Building successful online communities : evidence-based social design.* Cambridge, Mass. : MIT Press.

[31] V P Murty, I C Ballard, K E Macduffie, R M Krebs, and R A Adcock. 2013. Hippocampal networks habituate as novelty accumulates. *Learning & Memory* 20, 4 (March 2013), 229–235.

[32] Pierre-Yves Oudeyer. 2018. Computational Theories of Curiosity-Driven Learning . (June 2018), 1–24.

[33] Raymond Patterson, Ian Gellatly, Ofer Arazy, and Soobaek Jang. 2007. The effects of wikis characteristics on performance quality: The role of user motivations. In *WITS 2007 - Proceedings, 17th Annual Workshop on Information Technologies and Systems*. University of Alberta, 43–48.

[34] Charan Ranganath and Gregor Rainer. 2003. Neural mechanisms for detecting and remembering novel events. *Nat Rev Neurosci* 4, 3 (March 2003), 193–202.

[35] Paul J Silvia. 2006. *Exploring the Psychology of Interest.* Oxford University Press.

[36] R Simpson, K R Page, and D De Roure. 2014. Zooniverse: Observing the World's Largest Citizen Science platform. In *WWW'14*. ACM Press, New York, New York, USA, 1049–1054.

[37] C D Smock and B G Holt. 1962. Children's Reactions to Novelty. *Child Development* 33, 3 (1962), 631.

[38] James Sprinks, Jessica Wardlaw, Robert Houghton, Steven Bamford, and Jeremy Morley. 2017. Task Workflow Design and its impact on performance and volunteers' subjective preference in Virtual Citizen Science. *Journal of Human Computer Studies* 104 (April 2017), 50–63.

[39] Alexandra To, Safinah Ali, Geoff Kaufman, and Jessica Hammer. 2016. Integrating Curiosity and Uncertainty in Game Design. In *Proceedings of 1st International Joint Conference of DiGRA and FDG*. 1–16.

[40] Bianca C Wittmann, Nico Bunzeck, Raymond J Dolan, and Emrah Düzel. 2007. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage* 38, 1 (Oct. 2007), 194–202.

[41] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, and Xing Xie. 2014. Mining novelty-seeking trait across heterogeneous domains. In *WWW*. 373–384.

[42] Oliver Zuchowski, Daniel Schlagwein, and Kai Fischbach. 2016. “Open Calls” Rather than “Fixed Assignments”: A Longitudinal Field Study of the Nature and Consequences of Internal Crowdsourcing. In *Thirty Seventh International Conference on Information Systems*. 1–12.