Coordination in a Peer Production Platform: A study of Reddit's /r/Place experiment

by

Ben Armstrong

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2018

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Understanding the factors causing groups to engage in coordinating behaviour has been an active research area for decades. In this thesis, we study this problem using a novel dataset of crowd behaviour from an online experiment hosted by Reddit. This experiment allowed users to attempt to build an image alone, or to work collaboratively in the hope of building something greater. We use data provided by Reddit, in addition to crowd-sourced coordination information, in order to compare this experiment with a platform containing many similarities to our experiment: Wikipedia. Comparison with Wikipedia shows that many behavioural trends appear to generalize across domains. We go on to construct an agent-based model of the experiment, allowing investigation into the effects of spontaneous and planned coordination. We find that while coordinated work leads to significant productivity improvements in concentrated areas, there is little effect on the experiment as a whole as a result of coordination.

Acknowledgements

I would first like to think my advisor, Kate Larson, for her guidance and support. I am grateful to Kate for giving me the freedom to develop and pursue my own interests while always providing excellent advice and helpful discussion.

I would also like to thank Robin Cohen and Peter van Beek for agreeing to review this thesis.

Finally, I would like to thank my mother, Elda, for her unconditional support, Nancy and David for making me feel welcome in Waterloo, and Harold for encouraging my continued education.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Peer-production is a type of crowdsourcing that relies on primarily self-organized contributors to produce novel work. These platforms gain strength from widely decentralized and diverse users who are responsible for creating tremendous bodies of knowledge. Significant research has been done to understand the dynamics of these crowd workers, and the factors affecting productivity in peer production platforms.

Projects such as Wikipedia or Linux require the cooperation of thousands of strangers to generate high quality content. Learning the mechanisms that inspire individuals to participate in such peer production platforms could lead to massive improvements in productivity across a variety of domains including wikis, open-source software development, and academic research.

Peer production platforms are of interest for many reasons. Several researchers have studied how membership turnover affects the collective ability of a crowd to produce content [40, 51], while others uncover how group structure impacts the ability of a group to coordinate [26]. Other lines of research have considered how one's impressions of their fellow workers may affect their acceptance of others' contributions [31], and attempted to predict how long-lasting a user's membership in a group would be [50]. Typically, research in peer production identifies a particular component of the system, likely related to user behaviour, and aims to understand the effect that component has on some outcome of the platform, such as membership size.

In addition to our analytical study of a novel peer production platform we develop an agent-based model (ABM) which is used to simulate the platform we study. ABMs are tools used to simulate a vast range of topics, from infectious diseases [13], to international negotiation [15], and housing markets [19]. Structurally, an ABM is a model composed of a number of independent agents controlled by a number of parameters that are often calibrated to empirical data. Typically an ABM identifies important components of some real-world system and implements them, often in a simplified fashion, in a computer model. A simple example of an ABM is Axelrod's Tournament in which agents play a Prisoner's Dilemma game with other agents, each agent following a unique and independently chosen strategy [1]. This then allows a modeler to adjust various parameters and simulate how a system might change as a result of such alterations [23].

In this thesis we introduce a collaborative online experiment conducted by Reddit in April 2017, referred to as the Place Experiment, or /r/Place, in which users collaboratively placed pixels on a globally shared canvas. The Place Experiment dataset has not, to the best of our knowledge, been studied prior to this work.

Our initial contribution is to examine how well /r/Place reflects trends identified in previous studies of peer production platforms, focusing on the large body of research that has utilized the highly accessible Wikipedia dataset. Based on this previous research, we make several hypotheses about the effects of various user behaviours in /r/Place and use traditional machine learning techniques to see whether Place Experiment contributors behave in the expected manner.

We also develop an agent-based model to simulate the users of /r/Place. This model simulates a number of different types of behaviour that were observed during the experiment. We use it to experiment with the proportions of each behaviour type in the simulation and shed light on the effects that coordination had on the outcome of the real-world experiment.

The remainder of this thesis is laid out as follows:

- **Chapter 2** introduces work related to this project. We survey the peer production literature and discuss the topics that are of interest in the field of peer production platforms. We also sample from the enormous body of literature pertaining to ABMs. In particular, we consider the ways in which empirical data is used to construct ABMs and the ways in which coordination amongst agents has been studied.

- **Chapter 3** discusses background information on Reddit and explains the details of the Place Experiment. Here we also introduce the specific data sets we make use of, and some basic statistics about users in the experiment.

- In **Chapter 4** we characterize the Place Experiment as a peer production platform and study the extent to which results from previous work are shown in this context. We use linear regression to test several hypotheses regarding trends that other researchers have reported while studying Wikipedia.

- **Chapter 5** constructs an agent-based model of /r/Place. After initially verifying the accuracy of the model we adjust levels of coordination and alter the mechanism agents use to decide how to behave in order to gain insight into the global effects of individual actions.
- **Chapter 6** concludes with a summary of the overarching findings of our experiments as well as discussion of possible future research directions.

Chapter 2

Related Work

This thesis approaches the analysis of /r/Place from two disparate perspectives. We begin by comparing the experiment to peer production platforms, most notably Wikipedia, and searching for commonalities or differences. Wikipedia was used for comparison primarily due to the wealth of previous research considering the website as a peer production platform. Our second approach creates an agent-based model of /r/Place in an attempt to understand how external coordination affected the outcome of the experiment. This chapter follows a similar structure, first introducing peer production and work that may share similarities to /r/Place and then presenting agent-based models that gave some inspiration for the direction taken in Chapter 5.

2.1 Peer Production

Peer production is a term coined by Benkler to refer to a new form of content production made possible by the digital age. It can be described as a "process by which many individuals ... contribute to a joint effort that effectively produces a unit of information or culture" [5]. The term has been used to describe a variety of tasks that are performed online including open source software development, educational resources, and Wikipedia [6]. A great deal of academic research has studied many aspects of peer production, including characterizations of why users join and leave tasks, how they coordinate their efforts, and the various types of users who contribute to online communities. A number of these topics have analogues in the Reddit /r/Place experiment described in Chapter 3.

2.1.1 Types of User

Members of online platforms can be classified in many different ways: by activity level, diversity, or even intent. A significant body of work has studied users with exceptionally high or low activity levels. For instance, many online platforms are populated by a large proportion of *lurkers* who contribute not at all or extremely rarely [16, 38], and the vast majority of Wikipedia editors who make only one edit during their first 24 hours do not return to the platform [37].

Panciera et al. [37] shows that over 99% of Wikipedia editors who make a single edit during their first day on the website do not go on to become even moderately significant contributors. They also show that activity level in the first days of an account is strongly predictive of an editor's future contribution level, demonstrating a significant difference between future "Wikipedians" and others. Meanwhile, Priedhorsky et al. [39] reveal that a mere 0.1% of Wikipedia editors are responsible for nearly half of the words that are read on Wikipedia. It has also been shown that by McInnis et al. that of crowdworkers asked to comment on a community forum, those that contributed fewer comments tended to be less socially involved or trusting [32].

The impact of a diverse set of actions has been found by Karumar et al. [25] who demonstrated that a user's first session on a website (specifically, movie rating and recommendation community MovieLens) can be strongly predictive of their retention rate. The more diverse a user's first session is the more likely they are to become a lasting member of the community. Diversity of team structure has been studied on GitHub, where it was shown that increased diversity of gender and team experience level was associated with a higher team work output [45].Users may also vary in their intentions; on a successful platform it seems likely that most users contribute useful content in good faith but due to the openness of the internet it is possible for users to intentionally vandalize a platform. This has been studied most notably on Wikipedia where "edit wars" involving repeated deletions and additions of content by two or more users are well-known. Kittur et al. [28] show that over 10% of Wikipedia users have reverted at least one contribution of another user. If even a fraction of these reverts are in bad faith, or are reverting a malicious contribution, there is clear evidence of a difference in intentions between users. The prevalence of vandalism in Wikipedia articles is also studied by Priedhorsky et al. [39] who find that while the probability of viewing a damaged article is very small, it was increasing at an exponential rate during their period of study[1].

[1]The authors note that the introduction of vandalism-repair bots shortly before the publication of the paper may have stopped this exponential growth.

2.1.2 Membership Turnover

A fundamental component of peer production platforms is that contributors can join or leave the platform at any time with little to no cost. This has been observed to happen very frequently and is referred to as turnover. Turnover has been studied on Wikipedia, as well as on WikiProjects which are open groups of Wikipedia editors devoted to curating specific topics and articles [51].

Traditionally, high rates of membership turnover have been considered a harmful drain of expertise on projects [11, 30]. This view is challenged in peer production platforms. Ransbotham & Kane [40] study the history of 2,065 featured Wikipedia articles, articles that undergo a process similar to peer-review and are judged to be accurate, complete, and well-written, with the hypothesis that editor turnover increases the chance of an article being featured and not being un-featured. They find their hypothesis to be accurate and suggest that turnover is useful because when editors stop improving an article on Wikipedia their knowledge does not leave with them but remains a valuable part of the article for future editors to build upon.

These results are echoed by Yu et al. [51] who show that while WikiProjects experience a loss of value from users departing the project (value measured in terms of number of edits to pages under the project's purview) the increase of value from new users joining the project is greater resulting in a net increase in the project's efficacy as a result of membership turnover.

Interestingly, despite being a useful force, turnover may not be directly related to the skill level of users. Balestra et al. [3] study the effect of "fun" as a motivation for editing Wikipedia and find that users who remain on the platform for a longer duration tend to be more strongly motivated by fun at early stages.

2.1.3 Personal Preferences

A common theme in peer production research is that of project or group membership. On most platforms, users either belong to explicitly organized groups such as GitHub projects, or are part of less formal groups such as "users who have edited the *Computer Science* page on Wikipedia." Looking beyond membership size and duration, there is much research focusing on the motivations users have for joining groups and how their preferences affect their actions.

Yu et al. [50] study two types of connection to a group amongst WikiProjects: identity-based attachment in which a user feels attached to the purpose of the group, and bonds-based attachment in which a user connects with other members of the group interpersonally. Both of these mechanisms affect the types of group a user may prefer to contribute to, as well as the frequency or quality of their contributions. They find that while identity-based and bonds-based attachment both relate to increased productivity, identity-based attachment has approximately twice the effect of bonds-based.

On GitHub, a popular code-sharing platform, a member of a project could be defined as a user that that has contributed code to the project. In order to contribute code to a project a user must have their code approved by an owner of the project. Marlow et al. [31] show that owner's personal judgment of the would-be contributor's profile can have an impact on whether the new code is accepted. These decisions, based on the personal preferences of another user's impressions, may not always be accurate and are shown to occasionally be based upon incorrect or irrelevant information.

2.1.4 Coordination Amongst Users

Coordination has been shown to have major effects on the success of projects. Most notably, pages on Wikipedia as well as projects on WikiProjects are subject to explicit coordination.

Mockus et al. [33] study two open source software projects and find that a challenge in open source development is coordination. In practice, they find that projects reduce coordination requirements by keeping teams working on high-coordination tasks as small as possible. This also has the effect of freeing more developers to work on low-coordination tasks.

Kittur, Lee & Kraut [26] study individual Wikipedia pages, differentiating between less complex tasks that require little coordination such as correcting grammar or spelling and high coordination tasks that involve entire articles. They find that low-coordination tasks are able to accommodate large numbers of editors with no negative effects while high coordination tasks do not receive any benefit, and occasionally experience damage, from larger numbers of editors.

WikiProjects aim to address Wikipedia's disparity in volume of popular culture content and "traditional encyclopedia" content. Each WikiProject allows users to explicitly coordinate on adding to or maintaining a particular topic on Wikipedia. Kittur, Pendleton, & Kraut [27] study whether WikiProjects are effective at significantly shifting their members' behaviour towards coordination-oriented tasks such as discussion with other editors. They find that joining a WikiProject causes a Wikipedian to slightly increase their activity level, focus more on the topic of the project, and have more discussions with other users on project based activities.

2.2 Agent-based Models

Agent-based Models have been used to simulate a vast range of topics; most notably for our purposes they have been extensively employed to study empirical data of human behaviour [9, 52]. Within this area, ABMs have been used to explore many subjects, often with a game-theoretic perspective, such as: the impact of subsidies on rooftop solar panel adoption [53], tax compliance [8], and the timing of retirement [2]. While there is a significant body of literature, there is little work directly related to our problem domain. In this section we focus on identifying the methods in which other literature develop ABMs based on empirical data.

While some estimates suggest that as low as 14% of studies incorporate parameter fitting to their ABM [44], there are nonetheless a variety of methods in which models are fit to real-world data. Typically the type of data in question and the expertise of the researchers involved both play a role in determining how models are calibrated.

A number of models have made use of regression techniques to fit model parameters. Zhang et al. [53] study the factors affecting rooftop solar adoption over time, with model parameters fit to empirical data using linear regression. When developing a model to study cooperation in a public goods game, Wunder et al. use regression to fit agent parameters to data collected from a laboratory experiment with human players.

Monte Carlo Sampling has been employed by Berger and Schreinemachers to sample from soil properties and socioeconomic surveys and construct agents and simulated landscapes [7]. This approach creates a model that provides a statistically robust recreation of empirical data.

Often a more direct model of parameter calibration is utilized. Bloomquist [8] constructs a model in which agents may choose whether or not to cheat on their taxes; several parameters in this model are calibrated by manual experimentation. This method is very simple but it also allows for the application of expert domain knowledge to the problem. Manual parameter calibration can also be done by a search of the parameter space. Evans and Kelley [17] use a similar method when calibrating their model of landcover change. The model is fit using a gradient descent search to iteratively modify parameter weights until the model is an acceptably close fit to the empirical data.

2.3 Coordination Amongst Users

Coordination can present itself in a number of ways. Often, in crowdsourcing platforms it is assumed that workers operate independently and without any interactions between workers. In fact, it is frequently the case that social networks composed to crowdworkers lead to various forms of coordination between large swathes of workers. Gray et al. conducted interviews showing that workers on crowdsourcing platforms such as Amazon's Mechanical Turk frequently invited other workers to tasks they found preferential [21]. These results are echoed by Gupta et al. who find, through interviews with Indian workers, evidence of social networks of communication and collaboration between workers [22]. Yin et al. take this line of research further and, through a large-scale investigation, find large social networks of workers helping each other to find and complete the best tasks [48]. This coordination between workers is not an inherent part of any crowdsourcing platform but nonetheless has come to play a major role within the platforms.

There is also a fair bit of coordination that can occur naturally within a system, actors may work in concert with others without even realizing that they are coordinating their actions. Prior research on spontaneous coordination in human systems appears sparse; however it has been observed in many biological systems. Groups of ants, for instance, converge upon the shortest path to a food source by combining the experience of all workers but without any explicit communication [20]. We see this in artificial systems as well; Beckers et al. [4] develop puck-gathering robots based on termite behaviour. The robots act completely independently of each other yet manage to combine dozens of individual pucks into a single pile, rather than a chaotic mess which was an entirely possible outcome.

While examples of emergent coordination in human behaviour appear to be lacking, there is strong evidence to suggest that it exists elsewhere, both in nature and in simulated humans. As well, recent research by Jun et al. [24] suggests that participants of online experiments with differing motivations will behave differently, suggesting that the topic of (particularly external) coordination in the /r/Place setting may have an effect on the experiment.

Chapter 3

Reddit Place Experiment

This chapter provides a broad overview of the data used throughout this thesis. It begins with an explanation of the experiment that is the focus of the thesis and proceeds to detail the data of which we have made use.

3.1 The Reddit Place Experiment

On April 1^{st}, 2017 the social news aggregation website Reddit launched a collaborative social experiment referred to as Place. The experiment began as a white canvas of 1000×1000 pixels. Every 5 minutes, each user who signed in to a Reddit account that had existed before the experiment began was able to select from any of 16 colours and place a single pixel on any of the 1,000,000 cells on the canvas, including cells which other users had previously filled. Users were able to see the x and y coordinates of each cell, and less obviously, the identity of the user who had previously placed a pixel in that cell (if such a user existed). Importantly, the canvas on which users placed pixels was globally shared; each user contributing to the experiment saw the same canvas as every other contributor at a given time.

This process continued for approximately 72 hours, with over 16 million pixels being placed by over 1,000,000 users. Over these 72 hours, many formal and informal partnerships formed to pursue thousands of artistic endeavours. Some users attempted to destroy existing work, while others contributed to constructing their national flags, university crests, favourite video game logos, and more. 320 cells were not placed in at all while one cell was placed in 37,214 times. A time-lapse of the experiment can be viewed on Youtube[1]. Images of the canvas at various times can be seen in Figure 3.1. Notably, the experiment had an unknown duration, while it was happening there was no announcement of when it would end.

3.2 Data Used

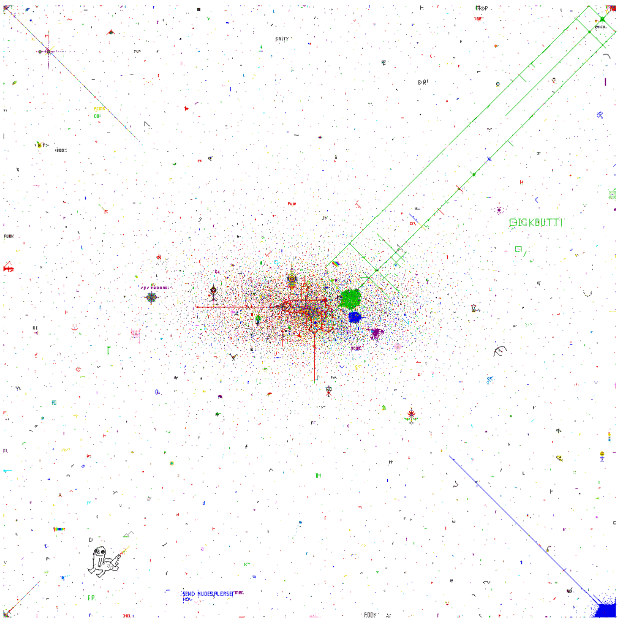
Our primary source of data is comprised of the entire list of pixels that were placed on the canvas. This dataset was released publicly by Reddit [14] and contains 16,559,897 records, where each record corresponds to a single pixel placement and contains 5 fields: *x_coordinate, y_coordinate, user, timestamp, color*. *user* is the hash of a Reddit account's username in order to preserve that user's privacy and *timestamp* is the Unix time at which the pixel was placed.

3.2.1 Place Atlas

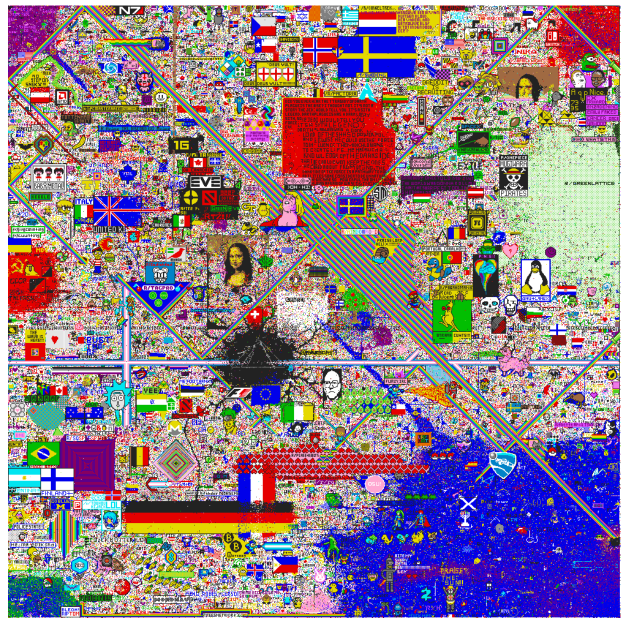
Our second source of data is provided via a crowdsourced collection of various areas of interest on the Place canvas, the Place Atlas. The Place Atlas was composed of both a forum on Reddit[2] and a separate webpage [41]. The forum on Reddit allowed any Reddit user to submit a JSON encoded entry to the Atlas, which contained several fields including a name and path outlining the entry on the canvas. We refer to each entry as a Region; each Region is a human-recognizable area that is visible on the canvas at the conclusion of the Place experiment. Several entries in the Place Atlas referred to Regions that were not visible on the final canvas, or were duplicates. After removing these, there remained 1389 Regions outlined on the final canvas. A number of these overlap, or are subsumed by larger Regions, which we allow to remain in the dataset.

We assume that each Region is an area that has some particular meaning or pattern that was recognized by most users who contributed pixels to the Region. Some of these Regions are quite complex and likely to have benefited from explicit planning, such as a recreation of the Mona Lisa, or a quote from *Star Wars: Episode III*. Others, like a spiral pattern or a mass of blue pixels, were much simpler and may have occurred through a more organic process. Many Regions can be identified quite clearly in Figure 3.1.

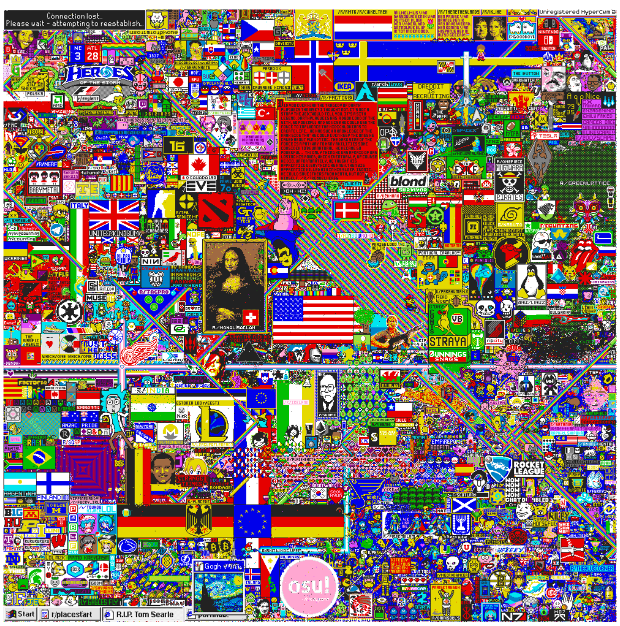
[1] https://youtu.be/XnRCZK3KjUY
[2] https://reddit.com/r/placeatlas



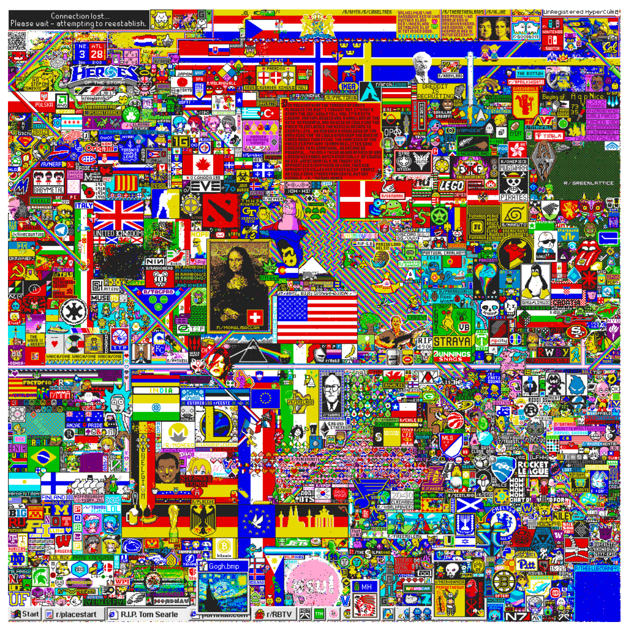
(a) 1 hour



(b) 24 hours



(c) 48 hours



(d) 72 hours

Figure 3.1: The Place canvas after 1, 24, 48, and 72 hours. The most dramatic changes appear in the first day but many changes can be observed even in the final 24 hours of the experiment.

	# Users	# Pixels	# Pixels Per User	
			Mean	Median
Inactive	616818	936189	1.52	1
Bot	2498	366269	146.62	115
Regular	547518	15256617	27.87	14

Table 3.1: Summary details about each user type.

When doing our analysis in Chapter 4 and Chapter 5, we treat the Regions labelled in the Place Atlas as a ground truth. We assume that all pixels within each Region are as they were intended to be by the majority of contributors. This assumption is certainly not perfect but we believe that (after having removed the Regions mentioned above) it is a reasonable approximation of the ground truth. Thus, throughout the thesis, we often refer to pixels as being *correct* or *incorrect* corresponding to whether or not they match the colour of the final canvas in the location they were placed.

3.2.2 Types of User

The 1,166,834 Reddit accounts that contributed at least one pixel to Place behaved in a variety of ways. Figure 3.2 shows that a very large proportion of users placed relatively few pixels while only a few users placed larger numbers of pixels. In order to allow analysis of different user behaviours we have placed each user into one of three categories: inactive, regular, and bot. These divisions are based on the number of pixels placed, and the frequency with which pixels were placed. A brief summary of each user type can be seen in Table 3.1.

Inactive Users

Inactive users are those users who placed less than or equal to the median number of pixels, 3 pixels. Inactive users are the most common type of user and represent those that did not invest a significant amount of time into contributing to the Place experiment. In total 616,818 inactive users placed a total of 936,189 pixels.

Bots

Many users employed scripts to help them build a design of their choosing, we refer to these users as *bots*. These scripts allowed a user to enter multiple reddit account credentials, and a source image. For each connected account, the script identifies a pixel on the canvas that does not match the supplied image and places a pixel matching the image using that account. This process repeats every 5 minutes. An example script can be found online on Github[3].

While it is impossible to know with certainty which users are making use of such scripts, we can make inferences based on the frequency and duration of a user's activity. We have two broad requirements for identifying bots:

1. Bots place pixels for a long time.
2. Bots place pixels very frequently.

Using these two requirements, we have classified a user as a bot if:

1. They are amongst the 100,000 most prolific users.
2. The mean period of time between their placing pixels, after removing outliers, was less than 350 seconds.
3. The standard deviation of the time between placing pixels, after removing outliers, was less than 200 seconds.

This process enabled us to identify bots even if the user had used a script sporadically, and placed pixels in a more human manner as well. This process identifies 2498 bots, which placed a total of 366,269 pixels.

Regular Users

Users who were neither Inactive Users nor Bots were classified as Regular Users. These users have placed more than 3 pixels, requiring at least 15 minutes between the beginning and end of their involvement with Place; they also did not place pixels extremely frequently for prolonged periods of time. These users, while not as common as Inactive users placed the vast majority of pixels. There were 547,518 Regular users placing a total of 15,256,617 pixels.

[3] https://github.com/Zequez/reddit-placebot

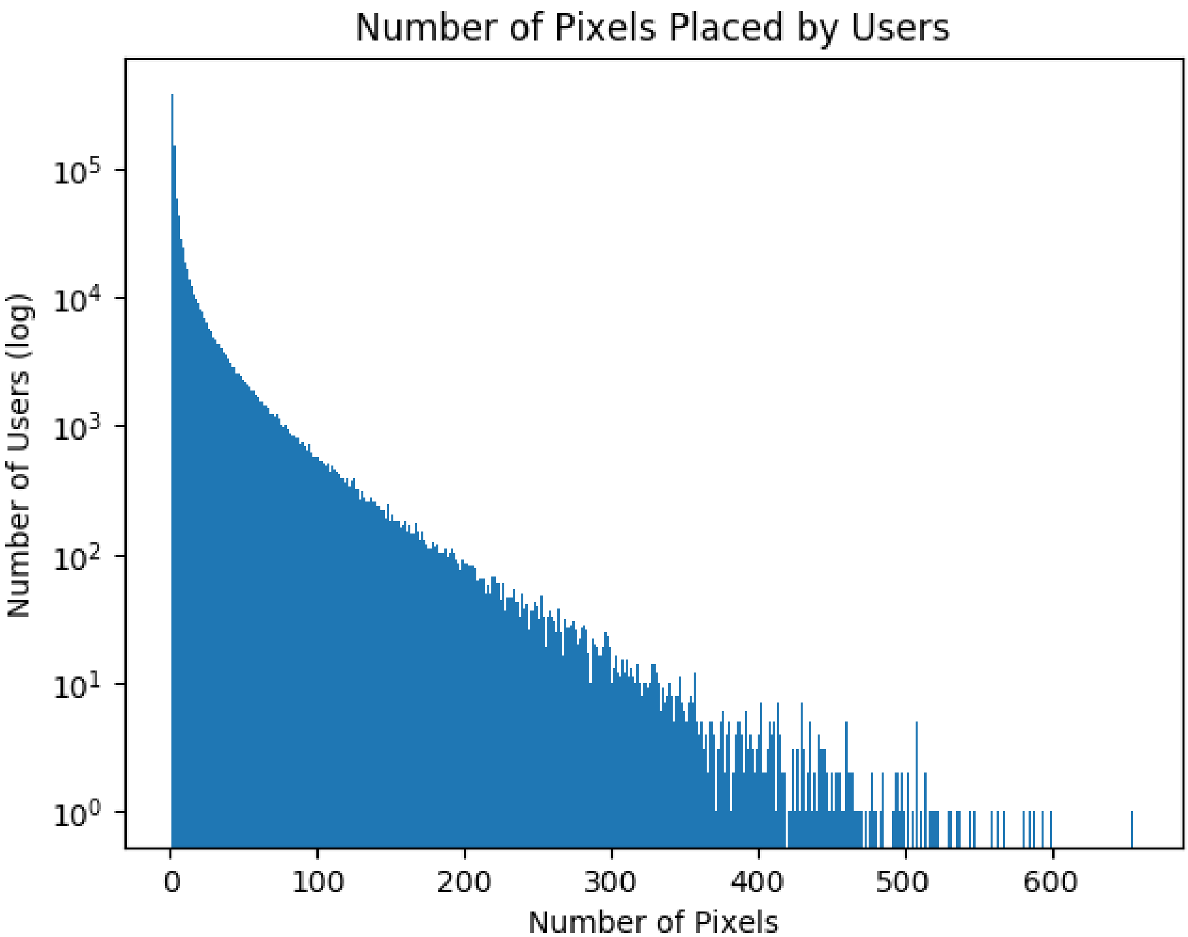


Figure 3.2: Base-10 log of the number of users placing each amount of pixels.

Chapter 4

/r/Place as a Peer Production Platform

The modern web has allowed for new forms of content production by enabling large numbers of individuals to work together. The massively wide-spread online encyclopedia, Wikipedia, is the most well-known example of such a platform. Github, an online source code sharing and version control platform is another. These platforms utilize what is known as *peer production* to facilitate massively large-scale content creation [6].

We therefore find it natural to question whether findings from other peer production domains transfer to behaviour observed in the Place Experiment as it contains all the essential elements of a peer production platform as described by Benkler [6]:

1. **Users have no centralized organization:** Reddit administrators gave users no instructions on how behave at the beginning of the Place Experiment or throughout its duration. Users were only given the suggestion that they create something.

2. **Users may have diverse motivations:** Since the announcement of /r/Place was visible to a significant and indiscriminately chosen portion of Reddit users, there is a virtual guarantee that users had a very diverse range of motivations for their contributions ranging from boredom, to wishing to be a part of a major worldwide event.

3. **Separation of governance from creations:** Again, there were no instructions given and no limits on what type of content could be created so all governance of the system was removed from the behaviour of contributors.

In particular, we focus on comparing and contrasting several general trends observed on Wikipedia to the Place Experiment. Wikipedia is used as a point of comparison for two primary reasons. First, there is a wide variety of literature studying Wikipedia as a peer production platform, more so than exists for any other platform. Second, /r/Place and Wikipedia are similar platforms in many respects. Users contribute to large projects, existing pages on Wikipedia or Regions in /r/Place, and most contributions are relatively insignificant but become important when viewed as a whole. There are notable structural differences as well (users on Wikipedia are able to see the entire history of the project they are contributing to, and may choose to make very significant changes) but we believe there are enough similarities to allow for an interesting comparison.

Our intention in this chapter is to identify which broad trends may be shared between both platforms and whether there is any common thread to the differences we identify. It may be that differences in how /r/Place and Wikipedia function are responsible for the majority of discrepancies we find.

4.1 Methods

In the following section we construct several hypotheses, largely based on prior research, and determine whether /r/Place supports them and in turn matches behaviour observed on Wikipedia. Specifically, we consider the following questions:

- Does membership turnover have any net positive effects on the development of Regions?
- Does a user who is loyal to a single, or small number of, Regions support the development of those Regions more than a user exhibiting no loyalty to any Region?
- Do more complex Regions develop differently than simpler Regions?
- Do different types of user affect the development of Regions differently?

Our primary tool to address these questions is linear regression. This technique easily allows a direct comparison between two datasets and provides results that are readily explainable. When considering the results of regression, we treat R^2 values as significant if they are greater than 0.3 as is generally considered acceptable in social sciences [51, 34]. Often hypotheses require taking observations of phenomena over a period of time; when this is the case, we sample based on one hour time periods. The Place Experiment lasted approximately 3 days so there are 72 time periods. Due to the large volume of data, p-values are reduced to nearly 0, thus we do not include them when presenting our results. As well, a base-2 log transform is applied to all data before regression, as this makes correlations more apparent in our data.

Our analysis typically follows a common pattern through several sub-sections. We begin by developing a hypothesis based on a component of previous research, then identifying one or more methods of testing that hypothesis, explaining the regression variables, and presenting our results along with a brief discussion.

4.2 Data Analysis

4.2.1 Member Turnover

Membership turnover occurs when a group exists across some period of time and some members depart while others may join. Generally, high levels of turnover are considered harmful to the group as it loses the expertise built up by long-time members[11, 30]. Yu et al. [51], and Ransbotham and Kane [40] both study the concept of member turnover on Wikipedia. They find that, contrary to previous work in other domains, membership turnover can provide useful benefits to a project via the addition of new ideas and the retention of previous members' work after their departure. A Region in /r/Place may not benefit from new ideas, since such ideas cannot be easily communicated and the design of a Region may be relatively inflexible. However, as Regions do retain the work of members once they depart and since new members may, at least initially, be more committed to a Region than older members we hypothesize that some level of membership turnover is useful for a Region.

Hypothesis 1 *Membership turnover tends to have positive effects for the growth of Regions.*

We test this hypothesis in two ways. First, we consider the direct effect of turnover by examining the impact of new and departing members on the correctness of contributions. Second, we examine the relationship between the average experience of Region contributors and the correctness of contributions to the Region. As described in Chapter 3, a contribution (a pixel) is *correct* if it matches the colour of the final canvas in the location it it placed.

	Pixels		Correct Pixels	
Independent Variable	Coefficient	R^2	Coefficient	R^2
Newcomers	1.03	0.89	1.00	0.68
Leavers	0.93	0.74	0.94	0.61

Table 4.1: Regression results showing the relationships between number of users joining and leaving a Region, and the number of pixels placed in Regions.

Turnover as change in membership

Yu et al. study the effects of both new and departing members on project productivity and find, somewhat unsurprisingly, that new project members are related to an increase in productivity while departing members are linked with a decrease in productivity. By noting that the association of new members with a productivity increase is stronger than that of departing members and productivity decrease, they suggest that membership turnover is beneficial [51].

We perform a similar test by identifying the relationships between the number of new members and departing members in each Region, and the amount of growth in each Region. Each variable is sampled across each Region and 1 hour time period.

Independent Variables

- **Newcomers:** The number of users contributing at least one pixel that did not contribute any pixels in the previous time period.
- **Leavers:** The number of users ho contributed at least one pixel in the previous time period and nothing in the current period.

Dependent Variables

- **Pixels:** The number of pixels placed during this time period.
- **Correct Pixels:** The number of correct pixels placed during this time period.

Our regression results, presented in Table 4.1, show that both newcomers and leavers are highly related to the number of pixels or correct pixels placed in a Region. This is likely a result of the large amount of turnover that Regions tend to experience. Figure 4.1 shows that in most Regions, a typical user will place between 1 and 3 pixels. Thus the number of users in a Region is strongly related to the number of pixels placed in the region, and

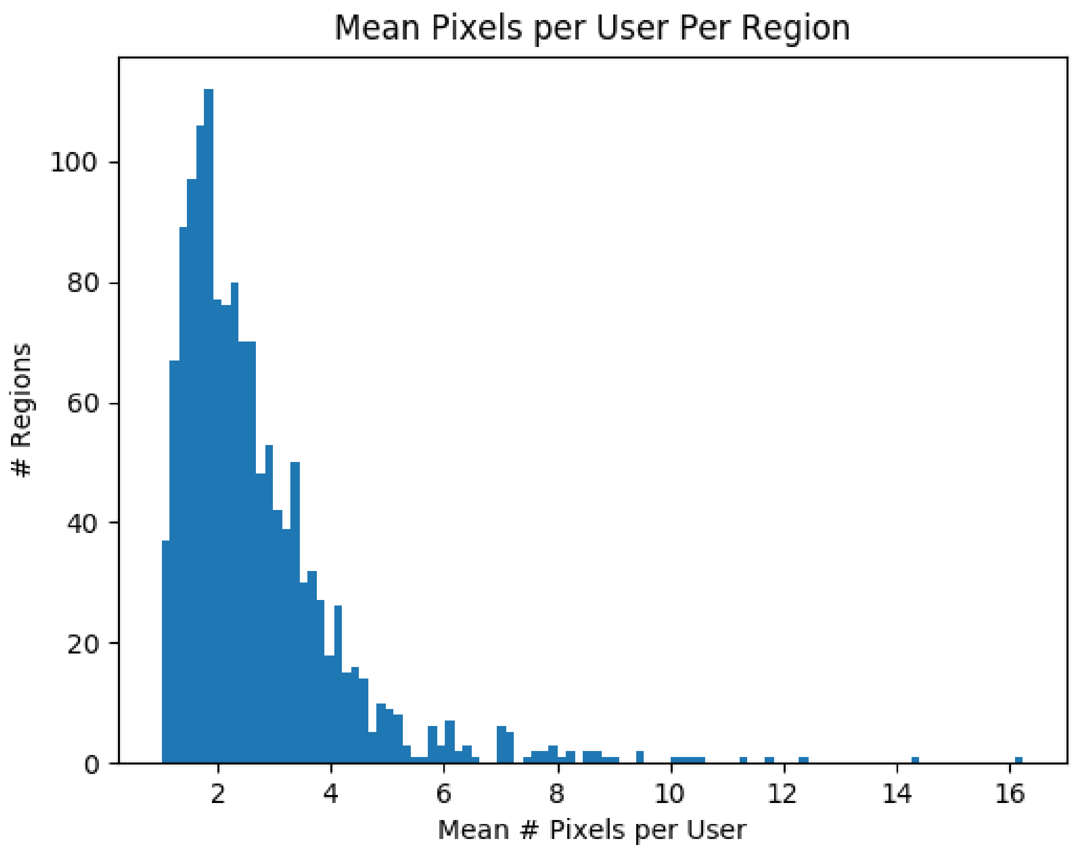


Figure 4.1: Average number of pixels placed by each user in each region.

if a large portion of users in a Region changes across each time step then the turnover in the Region will be strongly related to the number of pixels placed in that Region.

However, the coefficients in Table 4.1 show that newcomers have a slightly stronger relationship with the number of pixels placed in a Region than leavers. This suggests that newcomers tend to be slightly more productive than leavers, echoing the results of Yu et al. [51] and supporting Hypothesis 1.

Turnover as the opposite of experience

A second, less direct, approach to analysing the effects of turnover is inspired by Ransbotham and Kane, who suggest that average experience of users in a Region can be used as an opposite of turnover [40]. This approach found that projects with high levels of average experience were not as successful as projects with moderate levels of average experience. Thus, we consider the relationship between the average amount of experience users in a Region have, and the completion level of the Region in each 1 hour time period.

Independent Variables

- **Mean Experience:** The mean number of pixels each user has placed in the region before this time period.
- **Mean Correct Experience:** The mean number of correct pixels each user has placed in the region before this time period.
- **Mean Experience (no zeros):** The mean number of pixels each user has placed in the region before this time period, with values of zero removed. Particularly near a Region's start, smaller Regions are often composed entirely of users who are new to the Region. Removing values of zero allows for a clearer understanding of the role experience plays in Region development by removing the effect of users who placed only one pixel and focusing on the effect of users who had the opportunity to gain experience but may have chosen not to.
- **Mean Correct Experience (no zeros):** The mean number of correct pixels each user has placed in the region before this time period, with values of zero removed.

Dependent Variable

- **Region completion:** The number of correct pixels that are currently in the Region divided by the area of the Region.

Independent Variable	Coefficient	R^2
Mean Experience	0.45	0.18
Mean Correct Experience	0.49	0.22
Mean Experience (no zeros)	0.37	0.16
Mean Correct Experience (no zeros)	0.48	0.48

Table 4.2: Regression results between the completion level of a Region and the amount of previous experience of the users contributing to the Region.

Our results are summarized in Table 4.2, which shows a strong link between Mean Correct Experience (with zeros removed) and Region completion. In order to support our hypothesis, the highest levels of Region completion must correspond to Regions with moderate levels of average experience. Figure 4.2 shows that, in fact, the most complete Regions approximately correspond to the highest levels of experience. Thus, the hypothesis is not supported but neither is it directly refuted.

We have seen some support for this hypothesis when considering the direct effects of newly arriving and departing contributors to Regions, and a lack of support when considering the indirect effect turnover has on Region completion by means of average experience levels. Given that we have only positive evidence we will suggest that Hypothesis 1 is supported, though the effect is not extremely strong.

4.2.2 User Loyalty

Effect of contribution duration on loyalty

A well-studied phenomenon in psychology is the "bandwagon effect" in which people begin to do something purely as a result of other people doing it. That is, popular activities or people tend to become even more popular. This can play a role in the creation of viral videos [18] and formation of public opinion [36]. Often the feelings inspired by the bandwagon effect are less powerful than they would be if they had arisen otherwise. For instance, some sports fans may begin to support a team simply because that team is performing well but this support is unlikely to last once another team starts doing better.

This effect may be seen in /r/Place; users are likely to begin contributing to a Region once it begins to be recognizable. However, if this is their reason for contributing to the Region their support may not last. We suspect that users who have built the Region from

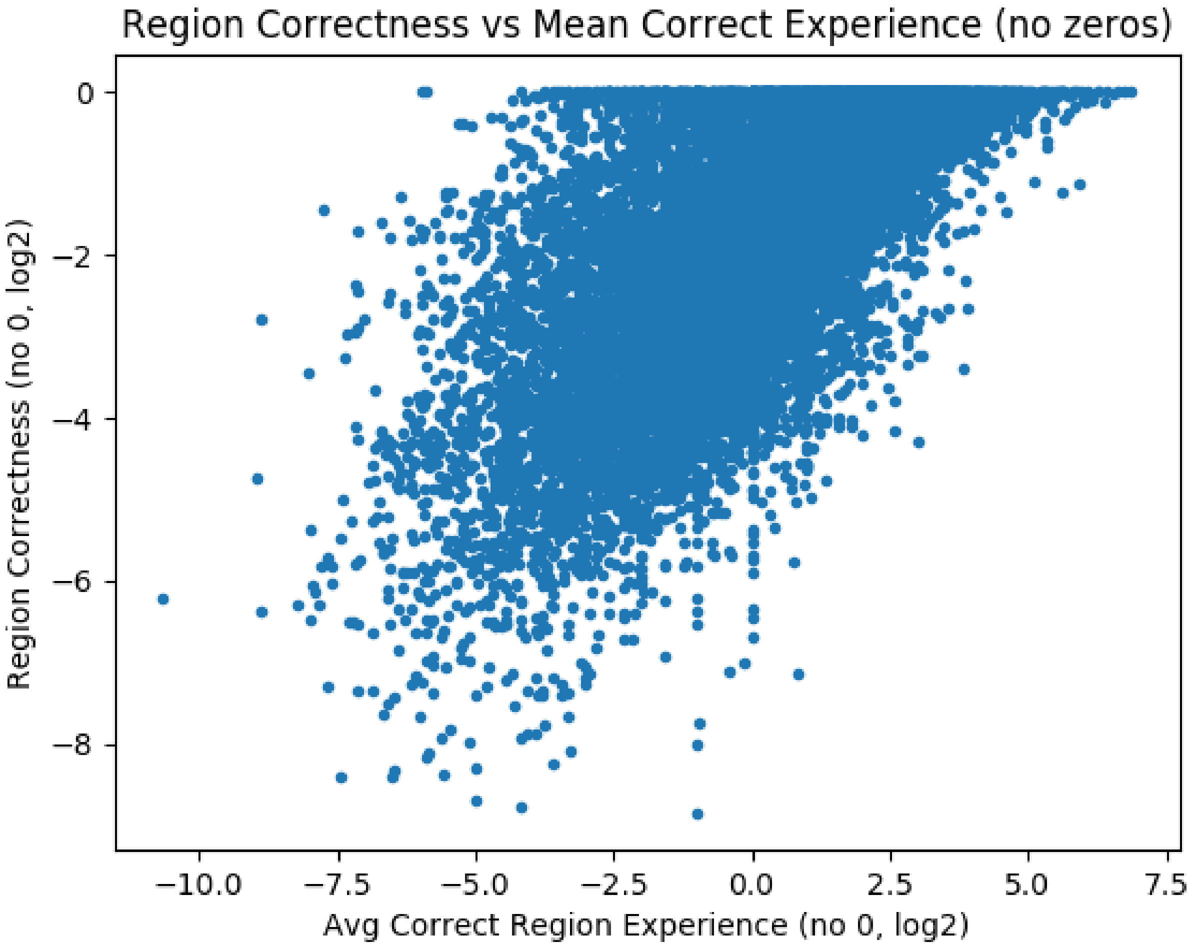


Figure 4.2: Average experience of users in Regions compared with the level of completion of Regions.

its beginning will show a more lasting support for the Region than users who join only after it is largely complete.

Hypothesis 2 *Users contributing to a region after it has experienced significant growth are less loyal to that region than users who contributed to the region before the growth.*

We test this hypothesis on several variables that may correspond to loyalty to identify any correlation with the level of completion of a Region. Each variable is evaluated across each individual region and one hour time period.

Independent Variables

- **Number of pixels:** The total number of pixels placed in the Region during the time period.
- **Number of correct pixels:** The total number of pixels placed in the Region that match the colour of the final canvas.
- **Number of users:** The number of unique users who placed at least one pixel in the Region.
- **Pixels per User:** The number of pixels divided by the number of users.
- **Mean of Number of Regions:** The mean of the number of Regions to which each user currently contributing to the Region contributed at least one pixel.
- **Mean of Previous Contributions:** The mean of the number of previous time periods in which each user currently contributing to the Region contributed at least one pixel to the Region.
- **Mean of Previous Good Contributions:** The mean of the number of previous time periods in which each user currently contributing to the Region contributed at least one correct pixel to the Region.

Dependent Variable

- **Region Completion:** The proportion of pixels placed in a Region that match the colour of that cell as it appears in the final canvas.

Of the above independent variables, the only one that is significantly correlated with Region Completion is Mean of Previous Good Contributions. Among the listed variables, this, along with Mean of Number of Regions, seems like the best analogue for user loyalty as it measure previous useful contributions to the Region. A larger value indicates the user has helped the group for longer periods of time. We can see in Table 4.3 how Mean of Previous Good Contributions is related to Region completion.

Data Used (Users contributing...)	Coefficient	R^2	Mean # Periods
All Data	0.55	0.29	2.53
...only before Region is 50% complete	0.21	0.03	0.45
...only after Region is 50% complete	0.05	0.06	2.22
...before and after Region is 25% complete	0.52	0.44	12.40
...before and after Region is 50% complete	0.50	0.41	11.70
...before and after Region is 75% complete	0.48	0.37	10.91

Table 4.3: Results of linear regression between the proportion of completeness of a Region and the mean number of time periods in which users contributing to a Region contributed to that Region. Users who contribute both before and after the Region experiences growth are much more closely related to the growth of the Region than other users.

We can see that there is no significant difference between users who contributed to a Region only while it was less than 50% complete or only while it was more than 50% complete. Users contributing before the Region is 50% complete are not more closely tied to Region success than those that first join the Region after it is 50% complete and, in fact, tend to contribute to the Region for fewer time periods. This may be due to the fact that the Region is not easily recognizable before it is mostly complete, so more users will tend to place random pixels in the Region, or because some Regions reach 50% completion early in the experiment and there is much more time after this milestone than before.

Further, we can also see in the final 3 rows of Table 4.3 that users contributing both before and after the Region passes certain threshold levels of completion are much more loyal than others. Part of this is due to the longer timeframe in which such users are able to contribute, however the difference is dramatic enough that an additional cause seems likely. It may be that these users are simply more committed to the Region than others – they have seen the Region experience significant growth and are more invested in its outcome than newer users.

Overall, we find that Hypothesis 2 is partially supported, albeit in an unexpected manner. While it appears that the most loyal users are those that contribute both before and after Regions experience significant growth, those that contribute after the Region is well developed are more loyal than early contributors despite their loyalty not being significantly correlated with growth in the Region.

Effect of number of commitments on Region growth

Our data shows us that users placing many pixels tend to contribute to multiple Regions but does not suggest whether some patterns of behaviour are more useful than others. We now consider the question of whether a user placing pixels in many different Regions is correlated with the proportion of pixels the user places correctly. Previously, Yu et al. have shown that on Wikipedia users joining and leaving a project are more useful if they are members of no other projects [51]. We suspect similar results will hold in our setting. Users placing pixels in multiple Regions are less likely to be strongly committed to the success of any particular Region, so we believe they will have a tendency to place fewer correct pixels in those Regions.

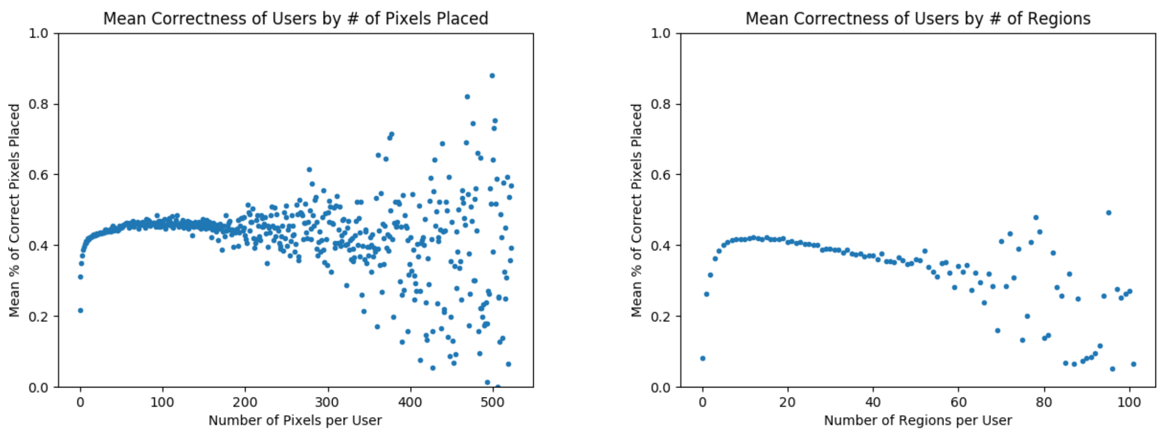
Hypothesis 3 *Users place a higher proportion of correct pixels in the Regions they contribute to if they contribute to fewer Regions overall.*

This hypothesis was tested by considering the relationship between both the total number of Regions a user contributes to and the total number of pixels they place (independent variables) with the proportion of pixels the user places correctly (dependent variable). We consider this proportion to be a reasonable approximation of how "useful" a user is. Table 4.4 summarizes our results.

Independent Variable	Coefficient	R^2
Total Pixels	-0.26	0.34
Total Regions	-0.39	0.33

Table 4.4: Linear regression results showing the effects of user loyalty on the proportion of correct pixels they placed.

Our regression analysis shows that the proportion of pixels a user places correctly is related with the total number of pixels a user places and the total number of Regions they place in. In each case, the relationship is negative, meaning that larger numbers of pixels placed or Regions placed within correlate with a smaller proportion of pixels placed correctly.



(a) (b)

Figure 4.3: Average correctness of users based on (a) the number of pixels they place and (b) the number of regions they contribute to. Correctness can be seen to decrease slightly more when users join an increasing number of Regions than when they place additional pixels.

This result indicates that as users place more pixels they are more likely to place an increasing proportion of incorrect pixels. Similarly, as users contribute to a larger number of Regions, they are more likely to place a larger proportion of incorrect pixels. However, as illustrated in Figure 4.3, the effect from the total number of Regions is stronger (Coefficient -0.39 for Regions, -0.26 for Pixels) indicating that placing in fewer Regions has a stronger effect on increasing a user's correctness than placing fewer pixels. Therefore, Hypothesis 3 is supported.

4.2.3 Team Structure and Task Coordination

The role of team structure in peer production is examined by Kittur et al. who find evidence that a large number of Wikipedia editors working on a low-coordination task is beneficial, whereas a large number of editors on a high-coordination task may be detrimental [26].

A reasonable analogue for task coordination in /r/Place may be the number of colours making up a Region, where the Region itself is the task. Certainly, a Region composed of only a small number of colours (such as the "Blue Corner" that can be seen at the beginning of an /r/Place time-lapse video[1]) is much more straightforward to contribute to than a complex Region composed of many different colours where it may be difficult to tell which colour a pixel is actually supposed to be (as in a Mona Lisa recreation, for instance). In simple tasks it is immediately obvious how a user might productively contribute, whereas in tasks requiring more coordination, contributing may require precise knowledge of the intended final result. This knowledge was often disseminated through subreddits, interest-targeted forums that make up Reddit, in the form of a bitmap image containing the grid of /r/Place filled in with exact colours in specific location showing what a desired final Region would look like.

Figure 4.4 shows the distribution of colours per Region which follows an approximately log-normal distribution. Nearly 90% of Regions contain between 2 and 7 colours with no Regions being composed of more than 14 colours (out of 16 possible colours).

Hypothesis 4 *Simple Regions composed of fewer colours experience more benefit from additional contributors than more complex Regions.*

We aim to test this hypothesis by considering whether there is any relationship between the complexity of a Region (as measured by our analogue: the number of colours in the Region) and any variable associated with the growth of a Region.

Independent Variables

- **Users:** Number of users contributing to the Region
- **Pixels:** Number of pixels placed in the Region
- **# Correct Pixels:** The total number of correct pixels placed in a Region.
- **% Correct Pixels:** The total number of correct pixels placed in a Region divided by the total number of pixels placed in the Region.
- **Pixels per User:** The total number of pixels placed in the Region divided by the total number of users contributing at least one pixel to the Region.

Dependent Variable

[1] https://youtu.be/XnRCZK3KjUY

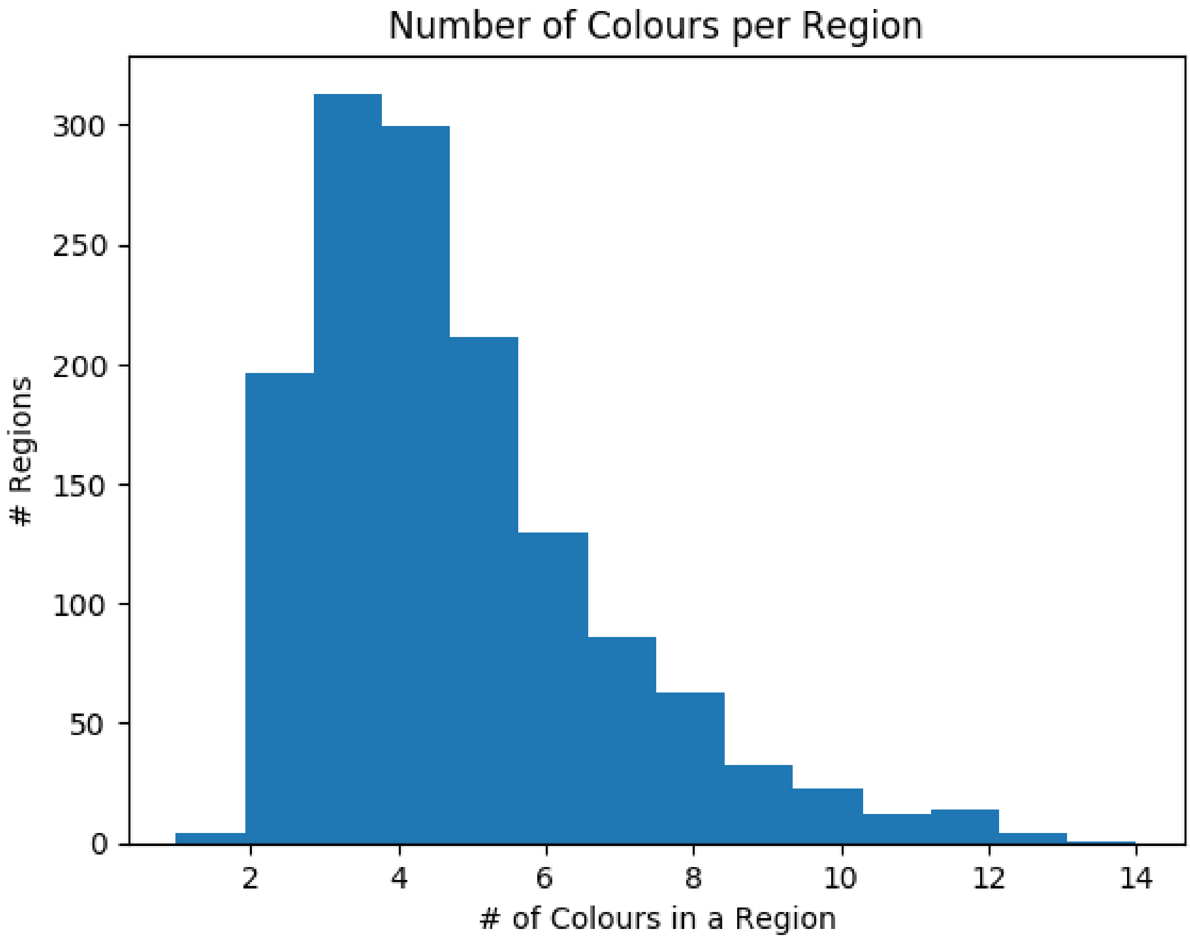


Figure 4.4: The minimal number of colours required to account for at least 95% of pixels in each of 1389 Regions at the end of the experiment. 82.7% of Regions were composed of between 2 and 6 (inclusive) colours.

Independent Variable	Coefficient	R^2
# Users	0.09	0.08
# Pixels	0.08	0.07
# Correct Pixels	0.06	0.05
% Correct Pixels	-0.18	0.02
Pixels per User	0.08	0.00

Table 4.5: Regression results showing (lack of) relationships between Region Complexity and several variables associated with the number of contributors to a Region.

- **Region Complexity:** The minimum number of colours to make up at least 95% of the pixels in the Region as it appears on the final canvas (95% is chosen by trial-and-error to generally capture nearly all of the desired colours in the Region while removing pixels that were placed erroneously and not removed before the end of the experiment).

Table 4.5 shows that we are generally unable to identify any relationship between our independent variables and the complexity of a Region. This may indicate that a Region's complexity has no bearing on the benefit it receives from additional contributors, or it may simply mean that our chosen proxy for Region complexity is inaccurate. Given the relatively low amount of complexity observed in practice in the majority of Regions on the /r/Place canvas both options seem possible. In either case, we find no evidence to support Hypothesis 4 and no clear proxy for complexity that would support further research is suggested.

4.2.4 User Type

In other domains, users have been divided based on their activity level, diversity of activity, duration, and intentions. In this experiment, there is only one action users can take which limits how much we can classify them. However, using the notion of a correct pixel and our previous categorizations of users based on quantity and frequency of activity we can begin to study whether our user categories correspond to fundamental differences between users.

We are primarily interested in whether there is a difference in the merit of each user type: Is a particular type of user more or less likely to place a correct pixel? Based on the idea that Inactive users are less committed and may simply be placing a pixel in an attempt to understand how the experiment works, we suspect they will tend to place fewer correct pixels. Similarly, bots are "motivated" with a very specific goal and might be expected to place almost entirely correct pixels.

Hypothesis 5 *Inactive users are more likely than Regular users to place an incorrect pixel.*

Overall, approximately 20% of pixels placed by Inactive users were correct while over 40% of pixels placed by Regular users matched the final canvas. Thus, the hypothesis is correct: Inactive users are more likely than Regular users to place an incorrect pixel. This can be seen further in Figure 4.5. The distribution of correct pixels placed by Inactive users is skewed very much to the side of incorrectness, while that of the Regular users is centered near 0.5, or half the pixels being correct.

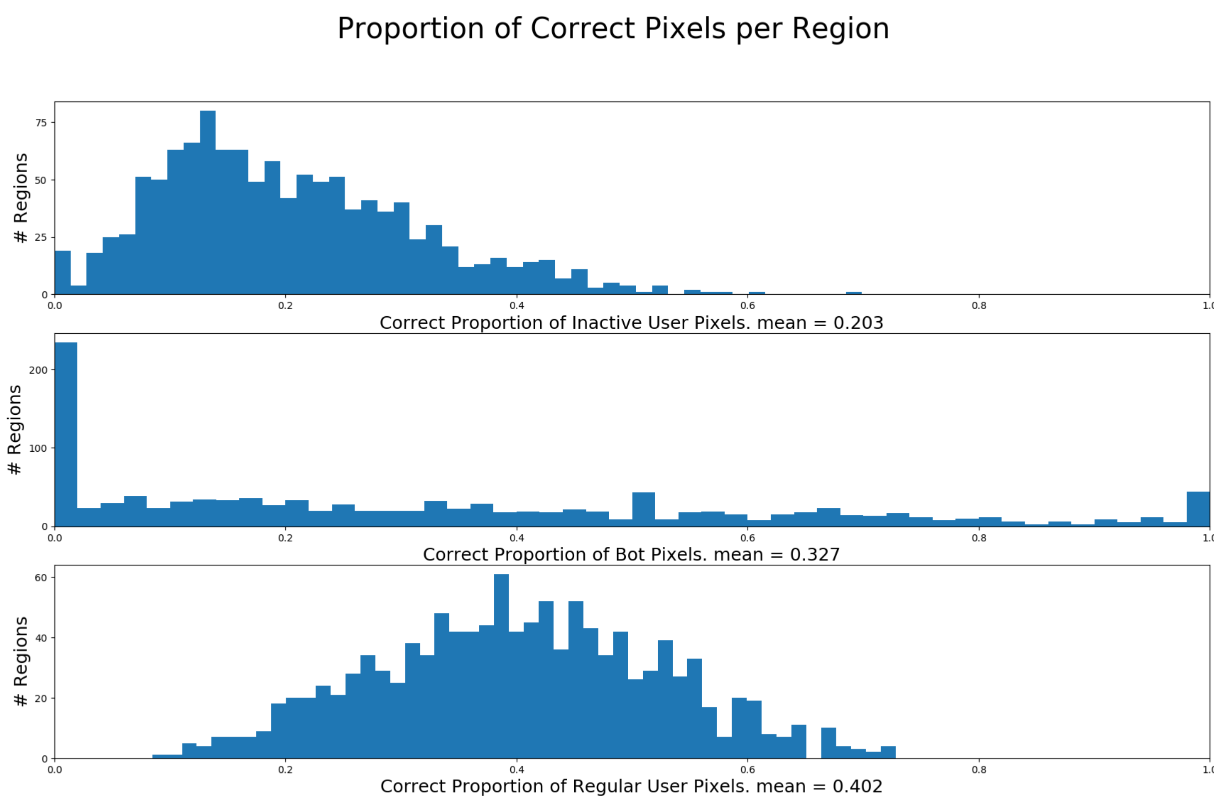


Figure 4.5: Proportion of pixels placed in each Region that match the final canvas, separated by user type.

Hypothesis	Description	Supported
1	Turnover increases Region growth.	Yes
2	Users are less loyal to completed Regions.	Partially
3	Users in fewer Regions are more helpful.	Yes
4	Simple Regions benefit more from additional contributors.	No
5	Inactive users place fewer correct pixels than Regular users.	Yes
6	Bots place more correct pixels than Regular users.	No

Table 4.6: Results of Section 4.2. When a hypothesis is supported, it matches the result observed in previous literature (except in the case of H5 and H6 which did not rely upon previous work).

Hypothesis 6 *Bots are more likely than Regular users to place a correct pixel.*

In the end, bots turn out to place proportionally fewer correct pixels than Regular users. However, the manner in which bots place pixels varies dramatically from the behaviour of Regular or Inactive users, lending credence to our classification of these users as non-human. On average, approximately 32% of pixels placed by bots matched the final canvas although, as seen in Figure 4.5, these pixels are distributed such that a similar number of regions existed at each level of correctness. This points to the possibility that many bots were set to build Regions that did not appear on the final canvas[2], and simply were not disabled once it became apparent to an intelligent observer that the Region would not succeed. These Regions would have shared a small, likely uniformly distributed, portion of pixels with each Region that did survive to the end, explaining the uniformity seen in Figure 4.5. The large number of Regions with no correct pixels from bots likely had only a very small number of pixels placed by bots, while the smaller spike in the number of Regions in which all bot pixels were correct would correspond with Region's to which bots intentionally contributed.

4.3 Discussion

Table 4.6 summarizes our hypotheses and the support for them. We have shown in Hypothesis 1 that new contributors to a Region tend to be slightly more productive than

[2]One set of four accounts for instance, drew several dinosaurs that were never visible on the canvas but by looking at only those user's contributions we can identify these Regions that were never seen during the experiment.

those leaving, similar to Wikiproject members studied by Yu et al. [51].

Hypothesis 2 demonstrates that while users contributing to a Region before it is popular may not be any more important to the Region than users joining after the Region is nearly complete it does show that users who are in the Region both before and after it is popular are disproportionately more useful than most contributors. Similarly, Hypothesis 3 shows that loyalty to a small number of Regions is more beneficial than widely divided loyalty.

Quite interestingly, the two hypotheses that are not supported at all are the ones least focused on typical users. When considering the effect of Region coordination requirements and complexity on Region growth in Hypothesis 4 we find that either colours make a poor proxy for complexity or there is no link between complexity and coordination. Similarly, we see the users suspected of being bots exhibiting behaviour very different than that of other users in Hypothesis 6. Neither of these hypotheses focus on the behaviours exhibited by the large majority of human users but rather on bots or structural components of the experiment.

Therefore, we suggest that behaviour exhibited by users during the Place Experiment is broadly similar to the behaviour observed across several research avenues on Wikipedia. This represents a very interesting generalization of human behaviour across entirely different domains. The differences between our results and the results of past work might be entirely explained by the differences between the Wikipedia setting and Reddit. We believe that /r/Place has potential as a new dataset which may be suitable for further research into peer production platforms.

Chapter 5

An Agent-Based Model of External Coordination

In this chapter we develop an agent-based model of the /r/Place experiment with the goal of investigating the results of coordination on the experiment outcome, as well as gaining insight into what mechanisms users may have used to decide upon their actions. Agent-based models are a tool widely used to model dynamic systems involving actions or interactions of agents [12]. An agent is a program that acts within an environment based on the state of the environment, the agent, and other agents. Agent-based models are composed of several agents, each of which models real-world actors such as individual humans [9] or even entire countries[15], as well as an environment which also represents an aspect of a real-world situation such as a housing market [19]. ABMs can provide an intuitive view of a situation while also allowing for exploration beyond what traditional analytical methods may allow [23]. In our model, agents represent individual /r/Place contributors and the environment upon which they act is the /r/Place canvas itself.

The chapter first describes the construction of our model, including a variety of different user types we have implemented as agents, then goes on to evaluate the strength of our model by comparing it with the actual experiment on several key metrics. Finally, we run a number of simulations in which we vary the level of coordination occurring in the simulation, as well as the method in which agents choose to which Regions to contribute pixels.

5.1 Model

5.1.1 Implementation

The simulation was developed using MASON, an Agent Based Modeling platform written in Java [29]. MASON is a discrete-time event driven simulator that provides a large tool-set to aid the development of agent-based models.

The two most important components of the simulation are the agents, and the model of the state of the world, or the environment. A world state is the core of any MASON simulation; we call ours the `Canvas`. It tracks information such as the current state of the simulated canvas and pixels, which agents currently exist, the Regions, and all of the parameter values.

Agents in MASON are created by implementing the `Steppable` class which requires creating a single method that can be run to perform the agent's action. This `step()` method contains all the behaviour of each agent and, in our case, performs two core actions: placing a pixel, and (possibly) scheduling the next pixel placement. Our agents are not scheduled to run in any regularly repeating loop but, instead, decide when to place their next pixel only after placing a pixel.

Throughout this chapter, the simulations we run are scaled in two ways for tractability. The canvas has a width and height of 500 rather than 1000, and each tick of the simulation simulates 10 seconds of real time so simulations last 25,920 ticks rather than 259,200 seconds. Regions used in the simulation are the same Regions contained within our cleaned-up version of the Place Atlas, except scaled to fit within the width and height of the simulated canvas. This was done to allow a more accurate comparison with the empirical data. As well, we retain the previous concept of correctness where a pixel is considered correct if it matches the colour of the final canvas at the location it is placed, though due to scaling a pixel placed on the simulated canvas at (x, y) is now correct if it matches the colour of the true experiment canvas at location $(2x, 2y)$.

5.1.2 Agents

Users expressed a wide variety of behaviours during the /r/Place experiment. Many contributed a few pixels then stopped, while others contributed hundreds of pixels over several days. Some users placed constructively, while others acted with the intention of destroying existing work. To attempt to capture some of the complexity seen in the experiment we implemented 5 types of agents based on real user behaviour.

Parameter	Description
Number of Pixels	The number of pixels the agent will place.
Correct Chance	The chance a pixel placed will be correct.
Number of Regions	The number of Regions to which an agent may contribute.
Chance of Covering Correct Pixel	The chance that a pixel will be placed above a correct pixel.
Time Between Pixels	The number of ticks between pixel placements. Scaled by simulation time factor. Always resampled until it corresponded to at least 300 seconds (the real minimum time between pixels)
Number of Agents per Tick	The number of agents that are created each simulation step. Scaled by simulation width and height factors.

Table 5.1: Parameters that control agent behaviour.

Each type of agent follows the same general life cycle of placing pixels that are either correct or incorrect with varying frequency until they have placed a pre-determined number of pixels, at which point they are removed from the simulation. More specifically, when an agent takes an action it:

1. Chooses a Region to contribute to.
2. Decides whether to place a correct pixel or not.
3. Decides on a location on the canvas within the selected Region to place the pixel.
4. Places the pixel.
5. Schedules the next pixel placement.

Agents have a variety of parameters governing their chance of placing correct or incorrect pixels, how many Regions they contribute to, how frequently they contribute, and how they are rewarded for their contributions. The value of each parameter, except for the *Chance of Covering Correct Pixel* parameter is fit to the empirical data for the users corresponding to that agent type. Those parameters that are distributions are chosen by fitting the corresponding data to each of Normal, Uniform, Beta, Exponential, and Gamma distributions and selecting the distribution which resulted in the lowest sum of squared errors. In nearly all cases this was a Beta distribution, when it was not we use the Beta distribution for simplicity (and the difference in sum of squared errors was minimal). Each agent has parameters corresponding to those shown in Table 5.1.

The five distinct types of agent are as follows:

Builders

Builders represent the users who actively tried to build well-formed Regions on the canvas. They follow the pure strategy of placing only correct pixels. A human user is classified a Builder if they:

1. Indicate a moderate level of commitment to the experiment by placing at least 12 pixels (which requires a minimum of 1 hour).
2. Place at least 90% of their pixels correctly. This shows a level of good intention not seen in many users.

There are 6383 users who meet the requirements and are considered Builders, this is approximately 0.5% of all users in the experiment. It may be that the actual fraction of users who made a significant fraction of correct contributions is larger than this, however as some Regions were covered by others the pixels of users who made the original Region may be considered incorrect despite being a positive contribution at their time of placement.

These agents are meant to represent users who tried to be as constructive as possible. Thus, whenever possible Builder parameters are set to the value that should lead to as great a benefit for the canvas as possible. Specifically, the *Correct Chance* and *Chance of Covering Correct Pixel* are set manually while other parameters correspond with values observed in the experimental data. Table 5.2 shows the default values given to Builder parameters in our simulations.

Saboteurs

Saboteurs are effectively the opposite of Builders. They represent the users who placed primarily incorrect pixels and attempted to sabotage the work of others. Saboteurs have similar requirements to Builders, they must:

[1] $Beta(a, b)$ indicates a value is sampled from a Beta distribution with parameters a and b.

Builder Parameters	
Parameter	**Value**
Number of Pixels	$12 + Beta(0.523, 154.4224)$[1] $\times 2060$
Correct Chance	1
Number of Regions	4 (median value from experiment data)
Chance of Covering Correct Pixel	0.01
Time Between Pixels	$Beta(0.6259, 418.353) \times 1124353/timeScaleFactor$
Number of Agents per Tick	$0.246/(widthScaleFactor \times heightScaleFactor)$

Table 5.2: Parameter values for Builder agents.

1. Indicate a moderate level of commitment to the experiment by placing at least 12 pixels (which requires a minimum of 1 hour).

2. Place at least 90% of their pixels incorrectly. This suggests a fairly strong intention of behaving poorly.

There are 27510 Saboteurs in the experimental data. The reason for there being so many more Saboteurs than Builders may be due to the previously mentioned phenomenon of pixels being placed with good intention and being labeled as incorrect. Again, *Correct Chance* and *Chance of Covering Correct Pixel* parameters are set manually to make Saboteurs highly destructive and the remainder of parameter values are calibrated to fit the empirical data. Saboteur parameter values are shown in Table 5.3.

Saboteur Parameters	
Parameter	**Value**
Number of Pixels	$12 + Beta(0.9068, 12687.7586) \times 404840$
Correct Chance	0
Number of Regions	9 (median value from experiment data)
Chance of Covering Correct Pixel	0.99
Time Between Pixels	$Beta(0.9087, 430.6039) \times 560774/timeScaleFactor$
Number of Agents per Tick	$1.059/(widthScaleFactor \times heightScaleFactor)$

Table 5.3: Parameter values for Saboteur agents.

Average Agents

The majority of users do not place as many pixels as Builders or Saboteurs, or place a variety of correct and incorrect pixels. A full third of users in the experiment placed only one pixel so Average Agents are, by far, the most populous agent, encompassing 1,133,031 users or 97% of users.

It is difficult to fully capture the dynamics of such a large variety of users, however the chosen parameter values do allow us to match several performance metrics between real and simulated data. Average Agents tend to place fewer pixels than other agents, and place approximately 16% of their pixels correctly. However, Average Agents are more likely to cover existing incorrect pixels than correct pixels when placing correct pixels, making them ultimately a force for overall correctness in the canvas. They place approximately 22% of their correct pixels above other correct pixels and 60% of their incorrect pixels above incorrect pixels. Table 5.4 shows each parameter value for Average Agents.

Average Agent Parameters	
Parameter	**Value**
Number of Pixels	$1 + Beta(0.5265, 50.903) \times 592$
Correct Chance	$0.1629 + 0.005 \times r$ r = proportion of correct cells in Region
Number of Regions	15 (median value from experiment data)
Chance of Covering Correct Pixel	$0.222 + 0.002887 \times h$ if pixel is correct $0.3 - 0.002778 \times h$ if pixel is incorrect h = # simulated hours since simulation start
Time Between Pixels	$Beta(0.4314, 655.6932) \times 416026/timeScaleFactor$
Number of Agents per Tick	$43.608/(widthScaleFactor \times heightScaleFactor)$

Table 5.4: Parameter values for Average Agents.

Spontaneous Coordinators

These agents model users who were very focused on a single Region, to the point where they almost certainly recognized the Region for what it was (ie. a flag or video game logo) and were intentionally assisting, or even beginning its development. They place pixels in only a single Region. This clear focus may allow them to be much more influential in the development of the Region than other users. Spontaneous Coordinators are similar to Builders but have one additional requirement, they:

- Contribute at least 12 pixels
- Place at least 90% of pixels their pixels correctly.
- Place all their pixels in a single Region.

These users were focused on a single Region but there is no evidence suggesting that they began contribution to /r/Place as a result of an external force. Thus, we treat Coordinators as organic, spontaneous users who decided to focus on a particular Region purely out of interest or a similar mechanism, as opposed to External Coordinators (described below) that we believe were influenced to contribute to a particular Region.

The most significant difference between Spontaneous Coordinators and previously described agents is that Spontaneous Coordinators are not created at some constant rate but instead are created only when a Region becomes coordinated. Each tick, every Region has a $\frac{0.006}{timeScaleFactor}$% chance of becoming coordinated if it has not previously been coordinated. This value has been chosen experimentally to be "close enough" to the empirical data. When this happens several Spontaneous Coordinators are created to contribute to that Region. In the real data, this may correspond to a Region becoming particularly noteworthy for some reason, or a Region related to a particularly devoted following. The value of each Spontaneous Coordinator parameter is shown in Table 5.5.

Spontaneous Coordinator Parameters	
Parameter	**Value**
Number of Pixels	$12 + Beta(0.5783, 51.5911) \times 1290$
Correct Chance	1
Number of Regions	1
Chance of Covering Correct Pixel	0.01
Time Between Pixels	$Beta(0.5064, 4829.8802) \times 17338405/timeScaleFactor$
Number of Agents per Region	$1 + Beta(0.5444, 596.286) \times 3624/(widthScaleFactor \times heightScaleFactor)$

Table 5.5: Parameter values for Spontaneous Coordinators.

External Coordinators

External Coordinators are meant to represent users who were directed to /r/Place from an external source that encouraged them to contribute to a specific region. These were typi-

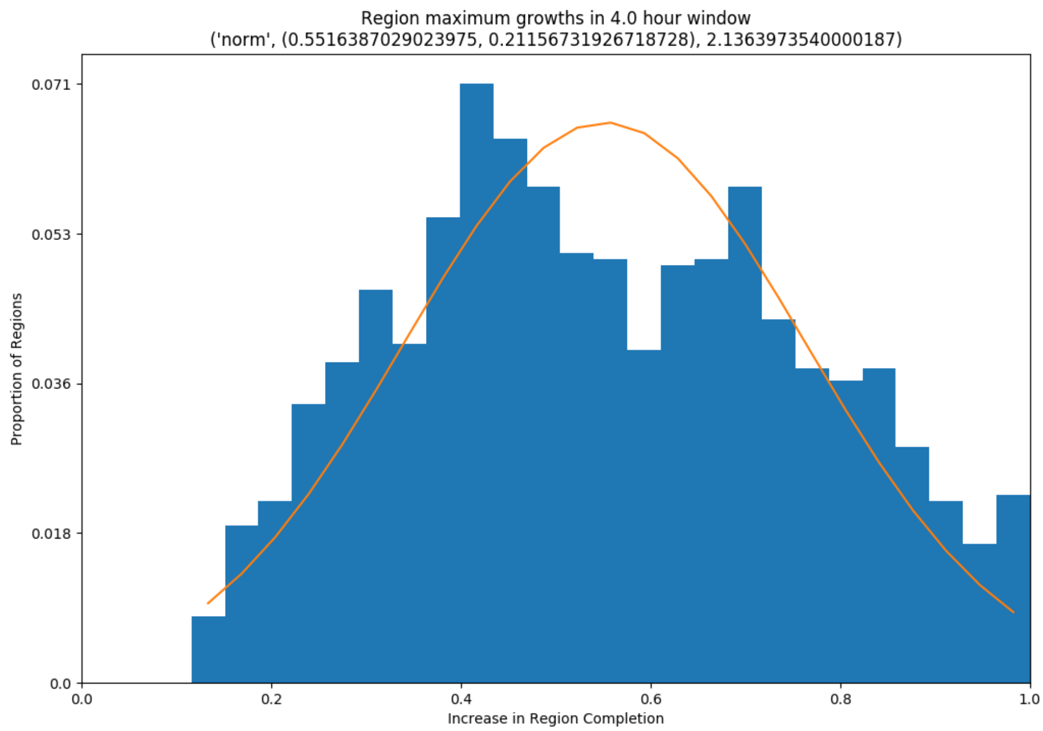


Figure 5.1: Histogram showing the maximum growth achieved in a 4 hour period by each proportion of Regions. A normal distribution with $\mu = 0.552, \sigma = 0.212$ is fit to the data with an error of only 2.14.

cally in the form of posts submitted to specific communities containing detailed schematics of an intended final Region[2]. Many Regions in the experiment experienced very rapid growth in short periods of time as seen in Figure 5.1. We hypothesize that popular posts outside of /r/Place may have contributed to the swift development of some Regions.

These agents are meant to replicate the rapid growth a Region may experience while an external post is highly visible to many Reddit users. As such, we have each External Coordinator represent the entire group of Reddit users who is attracted to the experiment from a post rather than using a single External Coordinator for each different user.

[2]Such as this post to a Star Wars focused community: https://www.reddit.com/r/PrequelMemes/comments/631283/rplace_the_darth_plagueis_project_megathread

Specifically, External Coordinators are found by the following process: We drew a histogram showing the maximum amount of growth each Region experienced in a fixed number of hours, for each of 1 through 23 hours. Each histogram was fit to a normal distribution; the histogram that best matched the distribution had a duration of 4 hours and can be seen in Figure 5.1. We are labeling Regions as being likely benefactors of external coordination if they experienced growth larger than one standard deviation above the mean growth in a 4 hour period. That is, a Region is externally coordinated if it increased its proportion of correct pixels by more than 71% in a 4 hour period. 260 out of 1389 Regions meet this definition. A subset of users contributing to externally coordinated Regions are labeled External Coordinators. In particular, External Coordinators are the users who meet the following two conditions:

- First pixel was placed in an externally coordinated Region.
- The majority of pixels placed in externally coordinated Regions were correct.

The parameter values seen in Table 5.6 used for External Coordinator agents reflect data that best fits the correct pixels placed by External Coordinators.

External Coordinator Parameters	
Parameter	**Value**
Number of Pixels	$1 + Beta(0.5373, 450.0514) \times 24091$
Correct Chance	1
Number of Regions	1
Chance of Covering Correct Pixel	0.01
Time Between Pixels	$14400/(NumberofPixels \times timeScaleFactor)$
Number of Agents per Region	1

Table 5.6: Parameter values for External Coordinators.

5.1.3 Region Selection

An important aspect of the simulation is how agents choose where to place a pixel. In our model, each pixel is explicitly placed in a Region at a semi-random location (with different chances of covering an existing correct or incorrect pixel as described above) and the Region is chosen randomly from a list of Regions selected for each agent when it is created. However, how that list of Regions for each agent is selected is a very important aspect of the model and can have a significant effect on Region development as well as the impact that coordination may have on the simulation.

There are a variety of methods we can use to select the Regions to which an agent might contribute. The simplest method is for each Region to be equally likely to be chosen at all times. This is unlikely to lead to accurate outcomes since, as we have already noted, Region growth rates vary greatly.

A second option is to assign each Region a random chance of being selected by an agent and for these chances to be randomized after some time. This would allow for Regions to have different growth rates but remains unlikely to match the behaviour of the real humans that were choosing where to place their pixels. Presumably real users actually gave some consideration to where they placed their pixels.

There may be some connection between the state of a Region when a user is considering where to place a pixel and the chance of that Region being selected. For instance, Regions experiencing high growth rates may attract more new users than other Regions in a sort of "snowball effect." This method may be accurate since users must have decided somehow where to contribute. The primary difficulty is in identifying which trait of a Region affects this decision.

A final method of choosing which Region to contribute to is to simply copy the experiment data. It is possible to identify the proportion of users contributing to each Region at each time window and use exactly that data in the simulation. This should be the most accurate way of selecting Regions. However, it reduces the chance that any interesting dynamics may emerge during the course of the simulation, and does not provide any insight as to the actual mechanism at play.

In our model we use the third method discussed above and hypothesize that there is a relationship between the proportion of pixels in a Region that are correct and the chance of that Region being selected. Specifically, we assume that the chance of a Region being selected is maximized when it is partially complete: that is, complete enough to be recognizable for what it is but incomplete enough that it clearly needs more work done.

We say that each Region has a chance of being selected equal to the chance of C being generated by a Normal distribution with mean 0.5 and standard deviation 0.25, where C is the proportion of pixels in a Region that are correct.

5.1.4 Data Collection

One of the challenges associated with this project is that of comparing data from the simulations with data from the experiment. The reasons for this are two-fold: First, the experiment has been analyzed in a Python project and we have developed a variety of tools to aid in this analysis while the simulation is being run in Java so the existing tools cannot easily be leveraged. Second, the simulation is being run in both a smaller space and smaller timeframe than the experiment. This means the resultant data will not have the same values as the experiment even if the simulation accurately recreates the experiment.

To solve these issues, we export a JSON-encoded file after each simulation containing various metrics that are tracked across the duration of the simulation. This file is used to construct a histogram or scatter plot of each metric based on the simulation data which is compared with a similar file generated based on the experiment data.

The file contains a dictionary with keys corresponding to each metric. If the metric is best represented by a histogram, each key points at a second dictionary containing 100 normalized bin values, and the width of each bin. Otherwise, the key points at a list of 432 values corresponding to the value of the metric at each simulated 10 minute interval (collected $\frac{600}{timeScaleFactor}$ ticks) for the duration of the experiment.


```
"results": {
    "Saboteur Ages": {
      "bin_width": 12.8,
      "bin_values": [ 0.268, 0.096, 0.042, ... ],
      "metric_type": "histogram"
    },
    "Number of New Builders": {
      "values": [3, 4, 4, 3, ...],
      "metric_type": "scatter"
    },
    ...
}
```


Figure 5.2: Sample of a simulation output file containing data to compare against experiment results.

A portion of a sample file can be seen in Figure 5.2. The file shows that 26.8% of Saboteurs exist for between 0 and 12 (inclusive) simulation ticks, 9.6% exist for between 13 and

25 simulation ticks, etc. Similarly, 3 Builders are created within the first 10 $\frac{600}{timeScaleFactor}$ ticks, 4 within the next $\frac{600}{timeScaleFactor}$, and so on.

5.2 Simulations

We perform simulations with a variety of parameter settings and collect several metrics to both validate our model and explore how changes in the model affect our comparison with the experiment data. We have three primary areas of interest when considering which simulations to run.

Model Validation

The first information we gather should focus on identifying the level of accuracy displayed by our model. This involves assessing the similarity between simulation and experiment for basic metrics such as the typical number of pixels placed by each user as well as less direct metrics such as the proportion of pixels in each Region that are correct across time.

When comparing a particular metric between actual and simulated data we use one of two methods to assess similarity, based on whether the data is in the form of a histogram or a series of single data points. For histograms we compare data using the cosine similarity [42], treating each list of bin values as one of the input vectors. The cosine similarity metric is bounded in the range $[0, 1]$, and larger values indicate greater similarity. The similarity value between two vectors A and B is defined as:

$$\frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

For series data we consider either the sum of differences between simulation and experiment data, or the average ratio between the series, deciding based upon whether the metric in question has been affected by one of the scaling parameters. As well, for both types of data we are able to visually inspect the similarity between the results.

Coordination

The primary goal of this chapter is to identify the importance of coordination, particularly external coordination. After validating the model, this is the clearest test to perform.

We can quite simply test the effects of coordination by disabling each different type of Coordinator agent and running simulations without them.

Due to our potentially noisy definitions of coordination, it may be that coordinators had a significant effect on the experiment outcome but no differences are noticed between the simulations with or without coordination. In this case, it is prudent to experiment with different levels of coordination. For instance, there may have been many users coordinating that did not meet our strict definitions of Coordinator or External Coordinator. This could be accommodated for by increasing the prevalence of each type of coordination in the simulation. For internal coordination, this is most easily done by simply increasing the number of Coordinator agents created in an internally coordinated Region. Since each External Coordinator models a period of coordination rather than an individual user, an External Coordinator is best strengthened by increasing the number of pixels it will place.

Region Selection

As described above, there are a variety of methods that can be used by agents to select which Regions they should contribute to. In the majority of our simulations we use the method of assigning each Region a probability of being chosen by an agent that corresponds to the chance of a Normal distribution ($\mu = 0.5, \sigma = 0.25$) generating that Region's proportion of correct pixels. While this method is able to generate acceptable results, comparing these results with those generated when using other Region selection methods may grant us additional insight into the accuracy of our chosen method. We compare four Region Selection methods, each of which is updated after every simulation tick (which corresponds to 10 seconds of real time):

- **Correctness-based**: The default method, as described above. This method was chosen with the idea that Regions that are semi-correct may be recognizable enough to attract new, interested users while Regions that are entirely correct do not express any need for new contributors and mostly incorrect Regions are not recognizable enough to attract attention.
- **Area-based Random**: Each Region has a chance of being selected proportional to a uniformly chosen random number scaled by the area of the Region. This method adds an element of chance to the process while also assuming that larger Regions are more likely to attract attention from new users.
- **Random**: All Regions are given a random chance of being selected, random values are chosen from a uniform distribution.

- **Real Popularities**: Each Region has a chance of being selected equal to the proportion of pixels that were placed in that Region during the real 10 minute window the simulation time corresponds to.

5.3 Results

In this section we will present the results of our simulations. First validating the model, then moving on to simulations studying the effects of coordination and the differences Region Selection methods cause in simulation outcomes.

5.3.1 Model Validation

Our first goal after developing our simulator was to verify that it is able to recreate behaviour from the experiment. We do this by defining several specific metrics of interest and comparing them between the simulation and experiment. A number of these metrics represent aspects of the experiment that are directly implemented in our model, such as the number of pixels each user should place. Data for other metrics, such as the time between a user placing their first and last pixels, is generated indirectly.

Direct Metrics

A summary of the results of metrics directly implemented in the model can be seen in Table 5.7, Table 5.8, and Table 5.9. We will now describe in additional detail how to interpret these results.

Table 5.7 shows the comparison between metrics best represented as a histogram. For each of these metrics, data was collected after every 600 real or simulated seconds for a total of 432 data collection periods over the 72 hour experiment. All data samples were then put into a normalized histogram containing 100 bins, and the cosine similarity between real and simulated data was computed. We also consider the width of bins in the histograms to understand whether the scale of the data is correct in our simulation.

In general, we have met with success when implementing the metrics in simulation. The Cosine Similarity measures shown in Table 5.7 are better when they approach 1 and a score of 0 indicates a lack of similarity between simulation and experiment metric data. We consider metrics fairly similar if they have a cosine similarity of approximately 0.7-0.8.

Bin Width refers to the width of each histogram bin after we have constructed a 100-bin histogram for each of simulation and experiment data. When it is more similar the metrics are more accurately implemented in the simulation. The scaling factors may affect bin width. However, as in the case of the Time Between Pixels metric where the ideal result would have Simulated Bin Widths as 10% of Experiment Bin Widths.

Metric Name	**Cosine Similarity**	**Simulated Bin Width**	**Experiment Bin Width**
Builder Pixels Per Agent (PPA)	0.914	1.78	3.64
Saboteur PPA	0.782	1.48	5.77
Average Agent PPA	0.998	3.63	6.55
Coordinator PPA	0.933	1.82	2.71
Builder Time Between Pixels (TBP)	0.638	14.46	2518.02
Saboteur TBP	0.503	13.27	2506.00
Average Agent TBP	0.870	44.65	422.14
Coordinator TBP	0.726	41.30	1904.65
Region Contributor Counts	0.994	962.67	6386.04
Correct Pixel Proportions	0.980	0.01	0.01

Table 5.7: Comparison of histogram metrics between simulations and experiment data.

Table 5.8 compares the number of users joining the experiment with the number of agents being created over time. Due to the scaling of both the width and height of the simulation canvas by half, a ratio of 4 is ideal and indicates that a proportionally correct number of users are contributing to the simulation. The values we see are sufficiently close to the ideal value that we feel the metric is being recreated accurately. A notable difference, however, is that in the simulation agents are created at a uniform rate. This is not the case with the experiment data which fluctuates with the day-night cycle, reaching a peak number of new users during the North American day times. Figure 5.3 shows a comparison of real and simulated Average user/Agent arrival rates over time. Plots for other agent

Metric Name	**Mean Ratio**
Number of New Builders	4.1
Number of New Saboteurs	4.02
Number of New Average Agents	4.01

Table 5.8: Ratio of the average number of users joining the experiment every 10 minutes to the average number of agents being created every simulated 10 minutes.

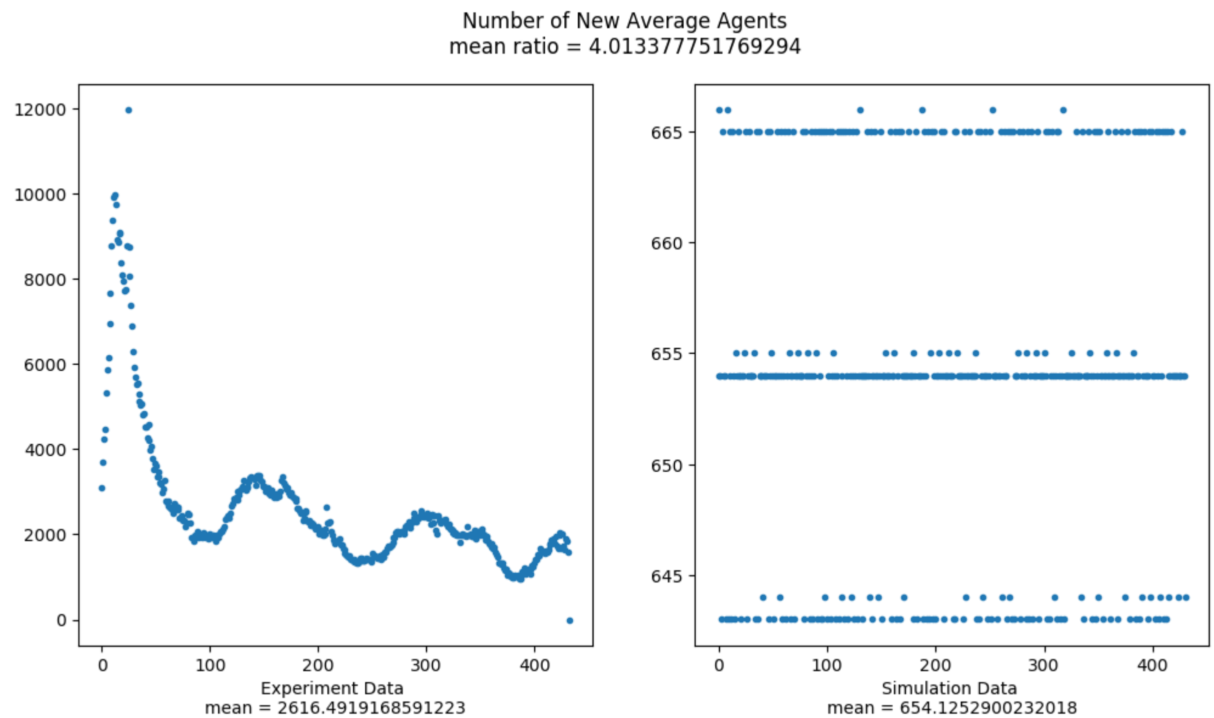


Figure 5.3: Rate at which new Average Agents, or their corresponding users, join the experiment or simulation across time.

types are quite similar. As there is no apparent change in behaviour caused by the number of users active at any time we do not expect our simulation's lack of a day-night cycle to affect any aspects of the simulation results.

The metrics shown in Table 5.9 contain data not directly controlled by the simulation parameters. During the simulation, the Number of Coordinated Regions and Number of Externally Coordinated Regions are generated by having a very small chance at each time for each Region to become Coordinated or Externally Coordinated. We can see in Figure 5.4 that the simulated rate of Regions becoming Coordinated follows a trend similar to the real data. While it is difficult to identify a cut-off for whether these metrics are similar or not we believe that they are similar enough for the purposes of the simulation. The reasons for this are two-fold: A similar trend is observed between real and simulated data when considering the number of Regions coordinated over time, and our analyses are not be affected by the specific number of Regions that are coordinated, but rather the ways in which those Regions develop. As well, the Fraction of Empty Cells is the proportion of

Metric Name	Mean Difference
Number of Coordinated Regions	36.95
Number of Externally Coordinated Regions	51.32
Fraction of Empty Cells	0.52

Table 5.9: Mean difference in value between real and simulated data for three metrics collected at each 10 minute interval.

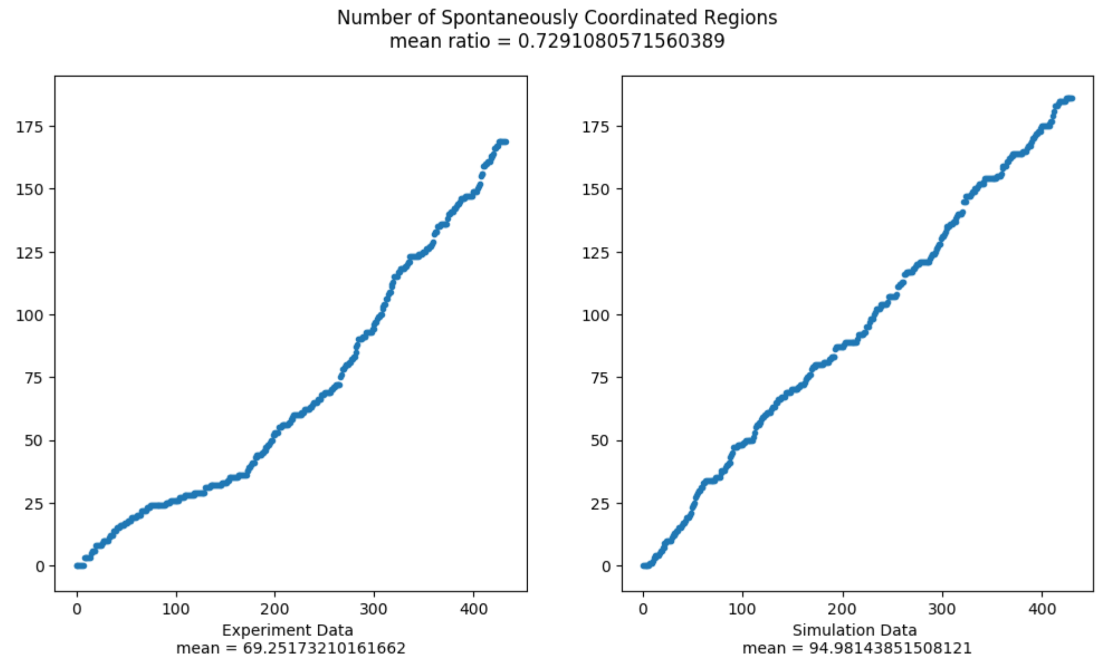
cells on the canvas that have not had any pixels placed in them. This metric measures the spread of pixels across the canvas and can give some understanding of how closely Region growth rates match the experiment data. As with the Number of Coordinated Regions, the value shown in Table 5.9 is much more meaningful presented visually, as in Figure 5.5. The figure shows that a sizable fraction of cells in the simulation are never covered throughout the entire simulation.

Indirect Metrics

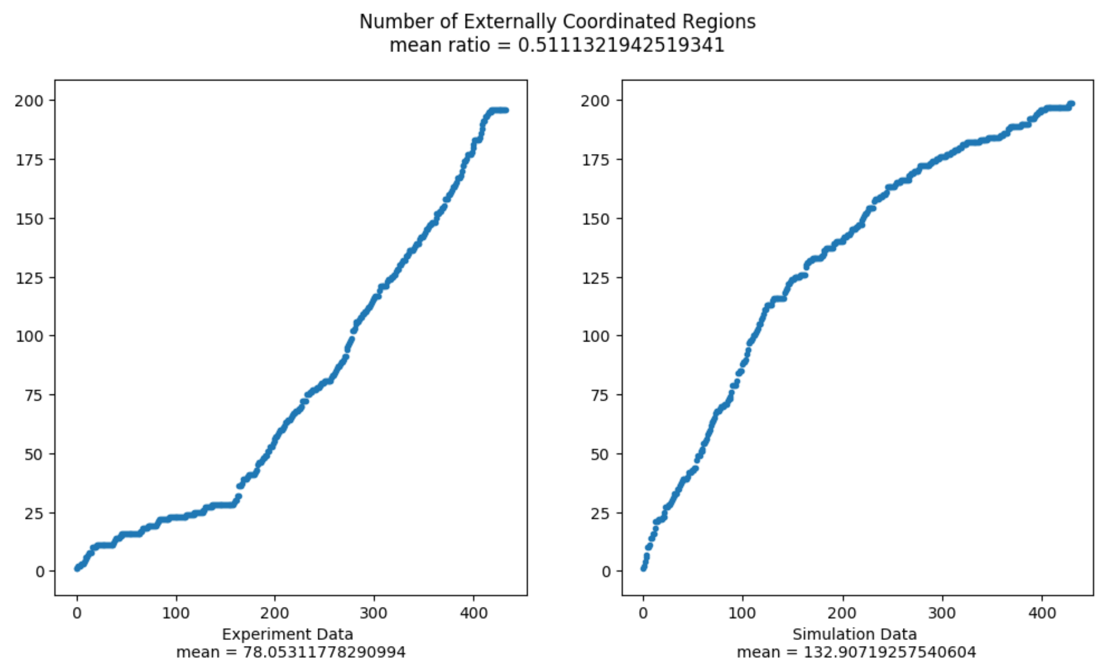
A number of metrics that we captured were not directly influenced by any one parameter. The metrics that we are more interested in are amongst these. Table 5.10 shows the indirect metrics that we collect. In general, the similarity between experiment and simulation amongst indirect metrics is slightly lower than the similarity between direct metrics but we still see values of at least 0.7, which is the bottom end of the range we considered to be minimally acceptable.

In particular, the age distribution of each agent type is relatively low. These metrics are a direct result of the distribution controlling time between pixels and the distribution controlling the number of pixels each agent places. This means that the error introduced by each of these distributions is magnified when looking at the age of each agent type. The other factor affecting these metrics is the relatively sparse amount of data for the Builder, Saboteur, and Coordinator agents.

The two metrics that we consider most interesting are those measuring Region development. Region Correct Proportions measures the proportion of cells in a Region covered by correct pixels. This is a collection of all 432 samples of each 1389 Regions, representing a view of the entire duration of the experiment/simulation. Figure 5.6 shows a comparison of the experiment and simulation. Both results are fairly similar. In particular, on the left side of each histogram we see that Regions are developed similarly when they are relatively incorrect. When approaching a state of full correctness, Regions in the simulation tend to



(a)



(b)

Figure 5.4: The number of Regions that are Spontaneously Coordinated, or Externally Coordinated, over time in experimental and simulated data.

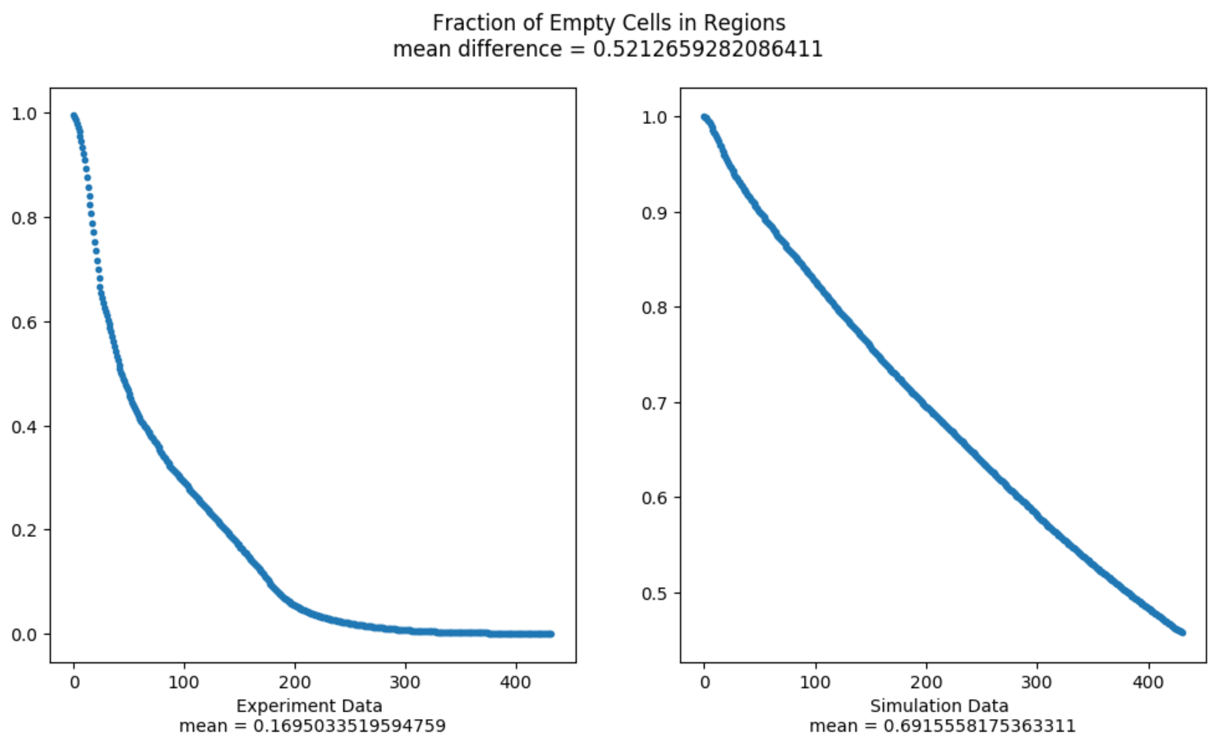


Figure 5.5: Fraction of cells that have not been placed at each 10 minute window of the experiment and simulation.

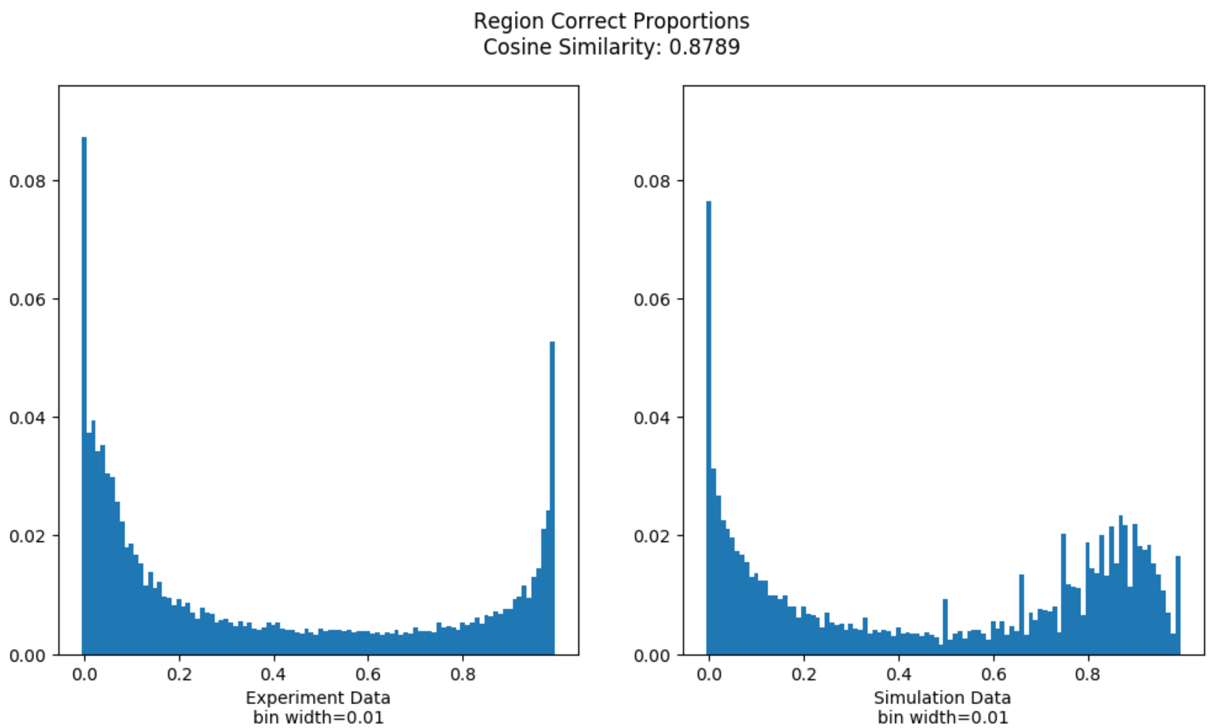


Figure 5.6: Fraction of cells that have not been placed at each 10 minute window of the experiment and simulation.

stop growing between 80% and 100% correctness while many Regions in the experiment reach near 100% correctness before the simulation ends (while all Regions are, by definition, entirely correct at the end of the simulation, the spike on the right side of the histogram in Figure 5.6 contains many more samples than those obtained in just the final sampling period). The lack of fully correct Regions in the simulation is difficult to explain, though it may be wholly due to a lack of intelligence in our simulated agents: the general behaviour trends are sufficient to mostly complete Regions, but some intelligence is required to get all the details fully correct.

We also look at how quickly Regions developed. The Region Maximum Growths metric shows, for each Region, the maximum proportion that Region increased in correctness over a (simulated) 4 hour period. This is the same definition we use when identifying external coordination. A number of parameters are likely to affect this metric, including the fraction of correct pixels placed by Average Agents, the chance of coordination/external coordination, the chance that Agents will cover correct pixels, the Region selection method

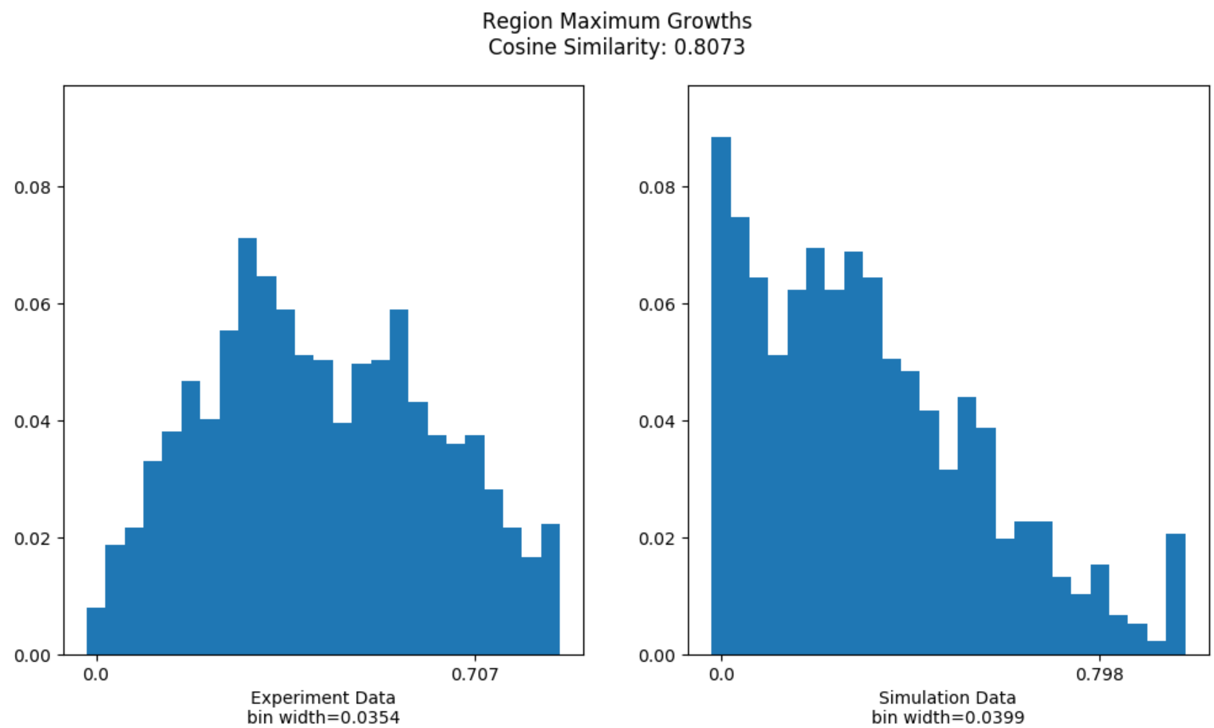


Figure 5.7: Fraction of cells that have not been placed at each 10 minute window of the experiment and simulation.

used, and perhaps others. Figure 5.7 compares the simulated and real data collected for this metric. The normal distribution in the experiment data is not fully represented. Instead, we see Region growth rates tend to be slightly lower in the simulation. However, the scale of both histograms is quite similar and if the Regions with the smallest growth rates were removed the data would be fairly similar. A reason behind the large proportion of Regions with lower growth rates may be the Region selection method used. This hypothesis will be explored later in this section.

5.3.2 Effects of Coordination

Our primary goal for our model is to investigate the effects that each type of Coordination may have had on the outcome of the experiment. We can fairly easily study this in our model by simply disabling Spontaneous Coordinators or External Coordinators and comparing the results with the results of our original simulation. As noted previously, our

Metric Name	Cosine Similarity	Simulated Bin Widths	Experiment Bin Widths
Builder Ages	0.709	443.58	4672.64
Saboteur Ages	0.706	373.32	5069.22
Average Agent Ages	0.855	511.16	4703.66
Coordinator Ages	0.708	178.04	3067.82
Region Correct Proportions	0.869	0.01	0.01
Region Max Growths	0.815	0.04	0.035

Table 5.10: Cosine Similarity of metrics that capture data not directly controlled by any single simulation parameter.

Metric Name	Cosine Similarity	Baseline Similarity
Region Correct Proportions	0.873	0.869
Region Maximum Growths	0.775	0.815

Table 5.11: Region development metric similarities when Spontaneous Coordination is disabled, compared with the cosine similarity of the same metrics using the default simulation parameters.

definitions of Coordination may be somewhat restrictive meaning that it may be informative to run simulations with larger amounts of Coordination in order to see whether it can have any interesting effects on the simulation. In this section we are primarily interested in the metrics measuring Region development: Region Maximum Growths, and Region Correct Proportions.

Spontaneous Coordination

We first study the effect that Spontaneous Coordinators have within our model. By simply having the model not create any Spontaneous Coordinators we are able to observe this. Table 5.11 shows the results of this comparison on the metrics we find most interesting (other metrics, excluding those collecting data directly related to Coordination, are not affected). We can see that there is no significant difference between these simulations and our baseline data.

However, we are also able to increase levels of Coordination to see whether it might have an impact on a simulation if it were more prevalent. Figure 5.8 (a) shows that when the number of Coordinators in each internally coordinated Region is increased by a factor of

10 there is not a significant change in similarity to the original experiment data. However, there is a noticeable increase in the proportion of Regions that are nearly entirely complete when compared with the default simulation data shown in Figure 5.9. This effect is to be expected; the Regions that are coordinated are having many more correct pixels placed within them, thus they are likely to be much more complete than previously. There seems to be little effect on the completion of other Regions; this is not guaranteed, as our default Region selection method makes highly correct Regions less likely to be contributed to by new agents.

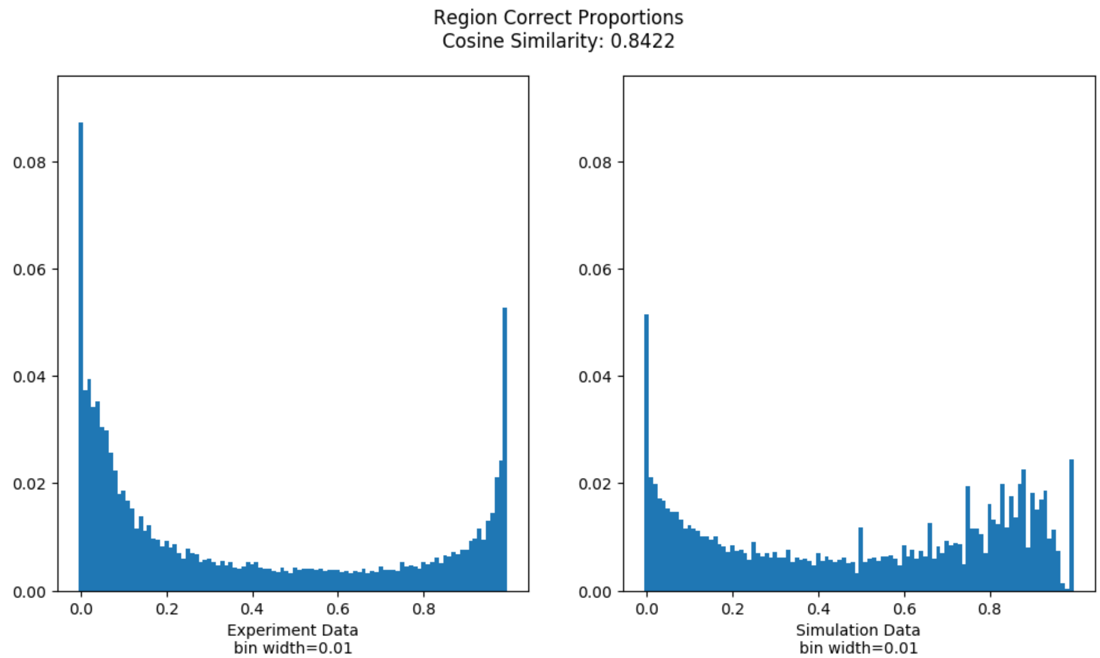
Interestingly, we observe that increasing the number of Spontaneous Coordinators in each Region seems to make Regions grow more slowly. A large portion of Regions shown in Figure 5.8 (b) grow a relatively small amount during their most rapid period of growth. The reasons for this are not totally clear but may be related to the Region selection method in use, making the coordinated Regions less likely to be contributed to by non-Coordinators.

External Coordination

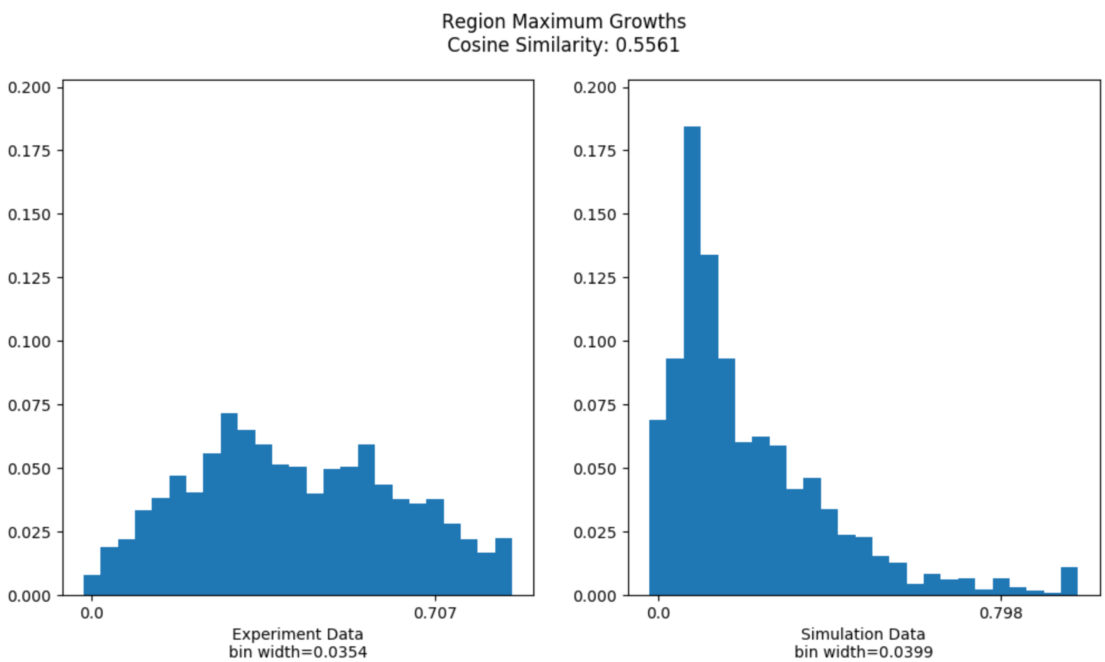
The second way in which Coordination is modeled in our simulations is by External Coordination, which aims to replicate the behaviour of users who came to the /r/Place experiment with the intention of contributing to a specific existing Region. These users were more numerous than users meeting our definition of Coordinators (users who contributed to only a Single Region) and therefore may have been more influential. We can study the effects of External Coordination in our model in much the same way as with Coordinators above: by entirely disabling it, or making it much more prevalent.

Again, we begin by running simulations with no External Coordination and see little change from the default simulation results. Figure 5.10 shows the results of a typical simulation with no External Coordination. There is a slight drop in the proportion of Regions that are nearly entirely correct, otherwise there is no noticeable difference from the baseline results shown in Figure 5.9.

When increasing the strength of External Coordination we find results similar to increased levels of Internal Coordination. Since External Coordination is modeled on a per-Region level rather than by individual agents, we cannot increase the number of coordinating agents in a Region. Instead, we increase the number of pixels placed during a Region's period of External Coordination by a factor of 10. The results, shown in Figure 5.11 are similar to the results of increased Internal Coordination levels. A small but unexpected difference, however, is that the increase in Region correctness levels is not



(a)



(b)

Figure 5.8: Region development metrics when using 10 times the default number of Spontaneous Coordinators in each coordinated Region.

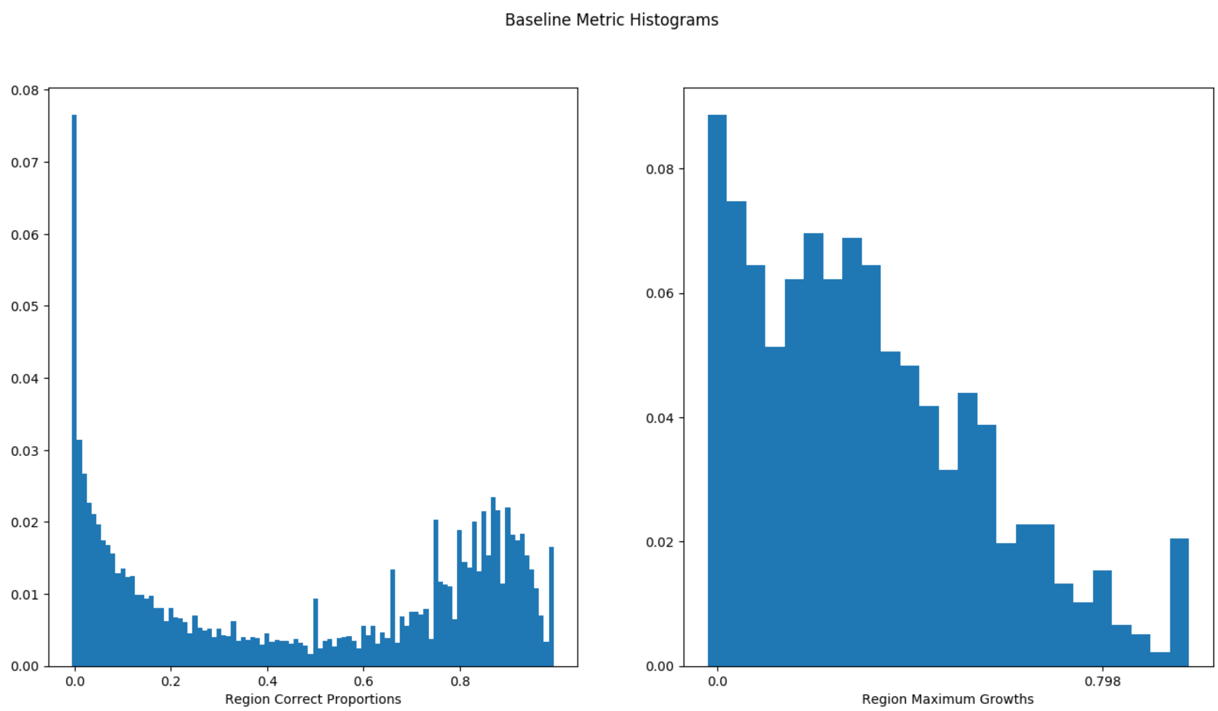
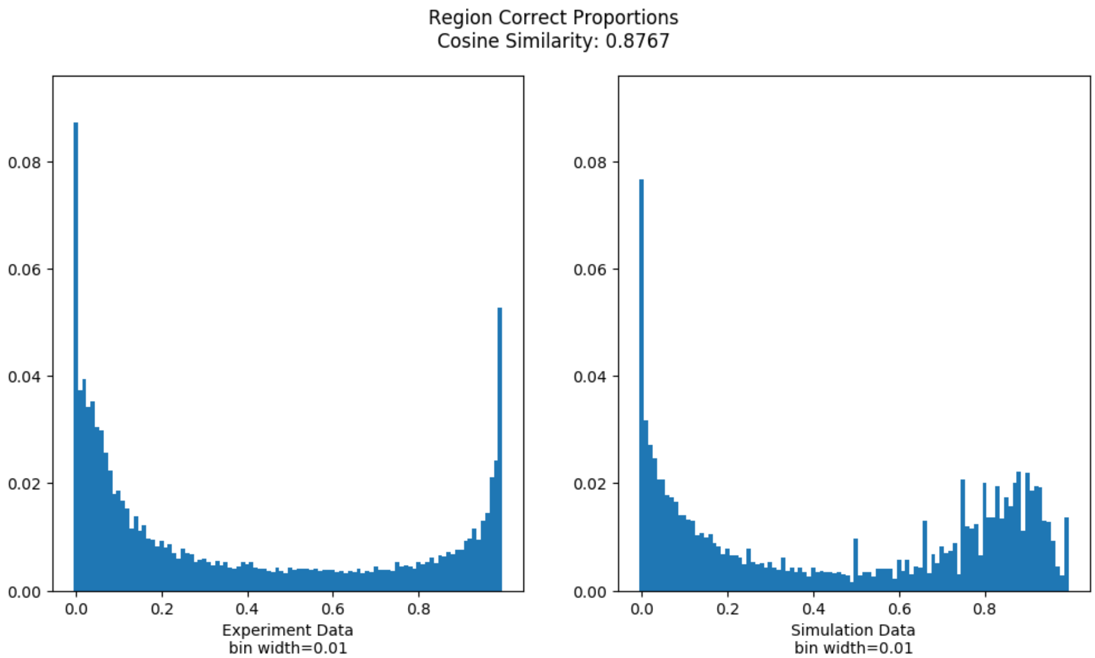
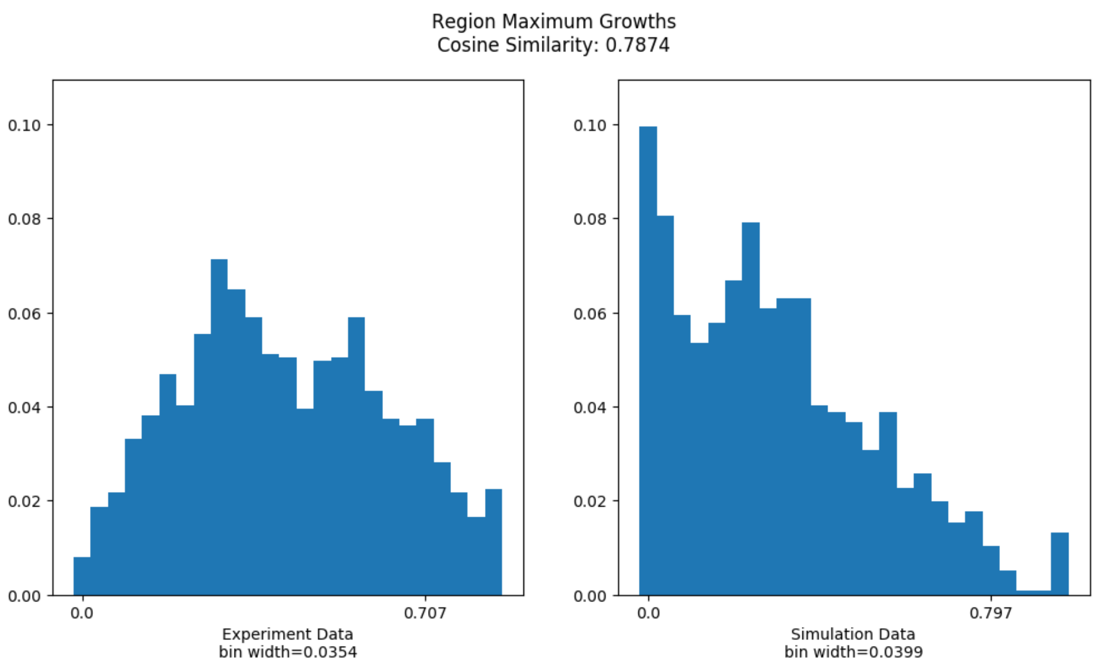


Figure 5.9: Region development metrics from simulations with default parameters.



(a)



(b)

Figure 5.10: Region development metrics when using no External Coordination.

limited to exclusively the right-most histogram bin. There is a spike in that bin, though several of the other largest bins also have a higher proportion of Regions in them.

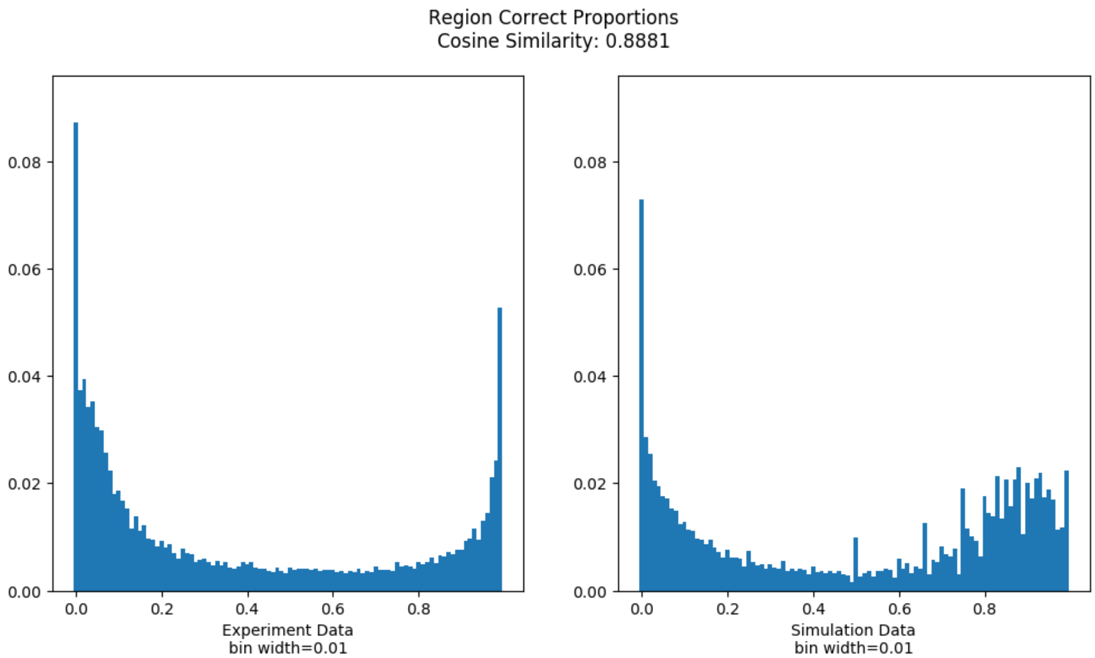
This departure from expectations is likely an artifact of how External Coordination is implemented in our model. When a Region begins to be Externally Coordinated, a number of pixels is chosen that will be placed to emulate pixels placed "by users arriving from an external source." These pixels are placed at a constant rate, spread evenly through a simulated 4 hour period. With Internal Coordination, a number of Coordinators are created when the Region begins to be coordinated and each Coordinator places their pixels independent of other Coordinators. This means that many more pixels are placed shortly after a Region becomes Coordinated, than when a Region becomes Externally Coordinated. This is reflected in Figure 5.8 and Figure 5.11: Externally Coordinated Regions tend to take longer than Coordinated Regions to become fully complete.

5.3.3 Region Selection

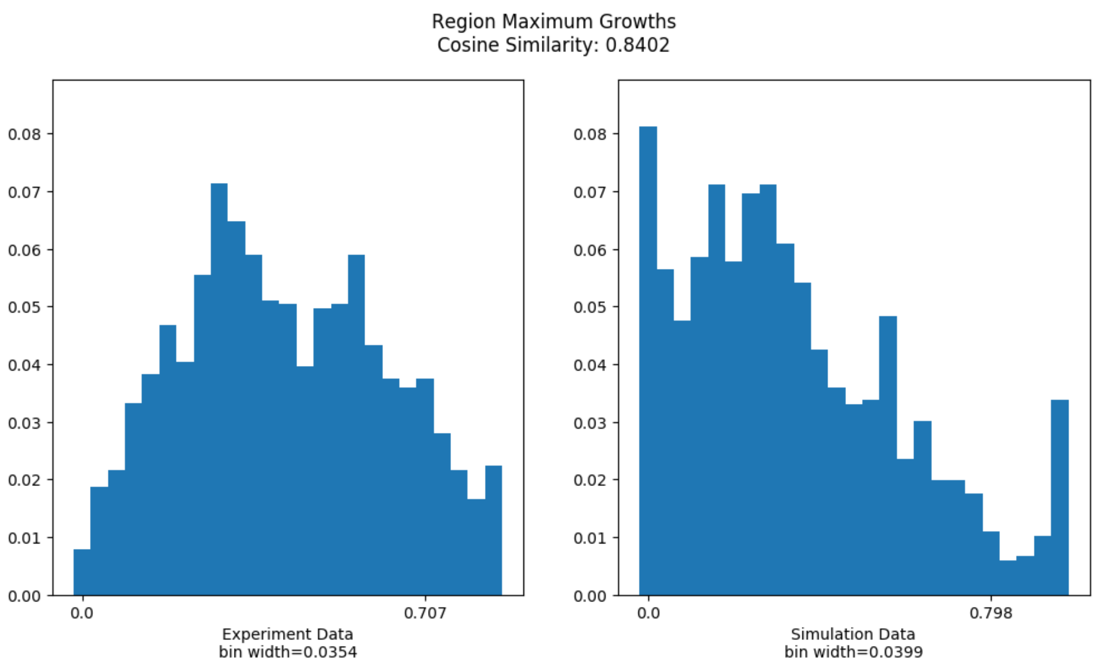
The Region Selection method controls how likely each Region is to be chosen as one of the Regions an agent may contribute to. There are a wide variety of possible methods to control this process, based on a number of different aspects of the model. We explore four such methods and compare their accuracy on the two Region development metrics we have used thus far. Section 5.2 introduces each of the methods we evaluate.

Overall, we find that our chosen Correctness-based selection method provides the most accurate results. This method follows a normal distribution and is most likely to select a Region that is 50% correct, while being minimally likely to select a Region that is 0% or 100% correct. While Table 5.12 shows that results with each method are not extremely different when measuring their Cosine Similarity, Figure 5.12 reveals that the differences are significant. Neither *Area-based Random* nor *Real Popularities* exhibit the same increase in correctness at the higher end of the scale that we generate with the *Correctness-based* or *Random* methods. Interestingly, results from the *Random* Region Selection method are quite similar to the carefully chosen *Correctness-based* method.

The results of the comparison become even more clear when considering the maximum growth achieved by Regions in each of the Selection methods. Again, *Area-based Random* and *Real Popularities* have a lower Cosine Similarity with experiment data while the *Correctness-based* and *Random* methods have noticeably higher Similarity. In this case, at least, the *Random* method does perform noticeably worse than the *Correctness-based* method. The same large spike in Maximum Growths can be seen in each of the worse



(a)



(b)

Figure 5.11: Region development metrics when External Coordination results in 10 times as many pixels as default.

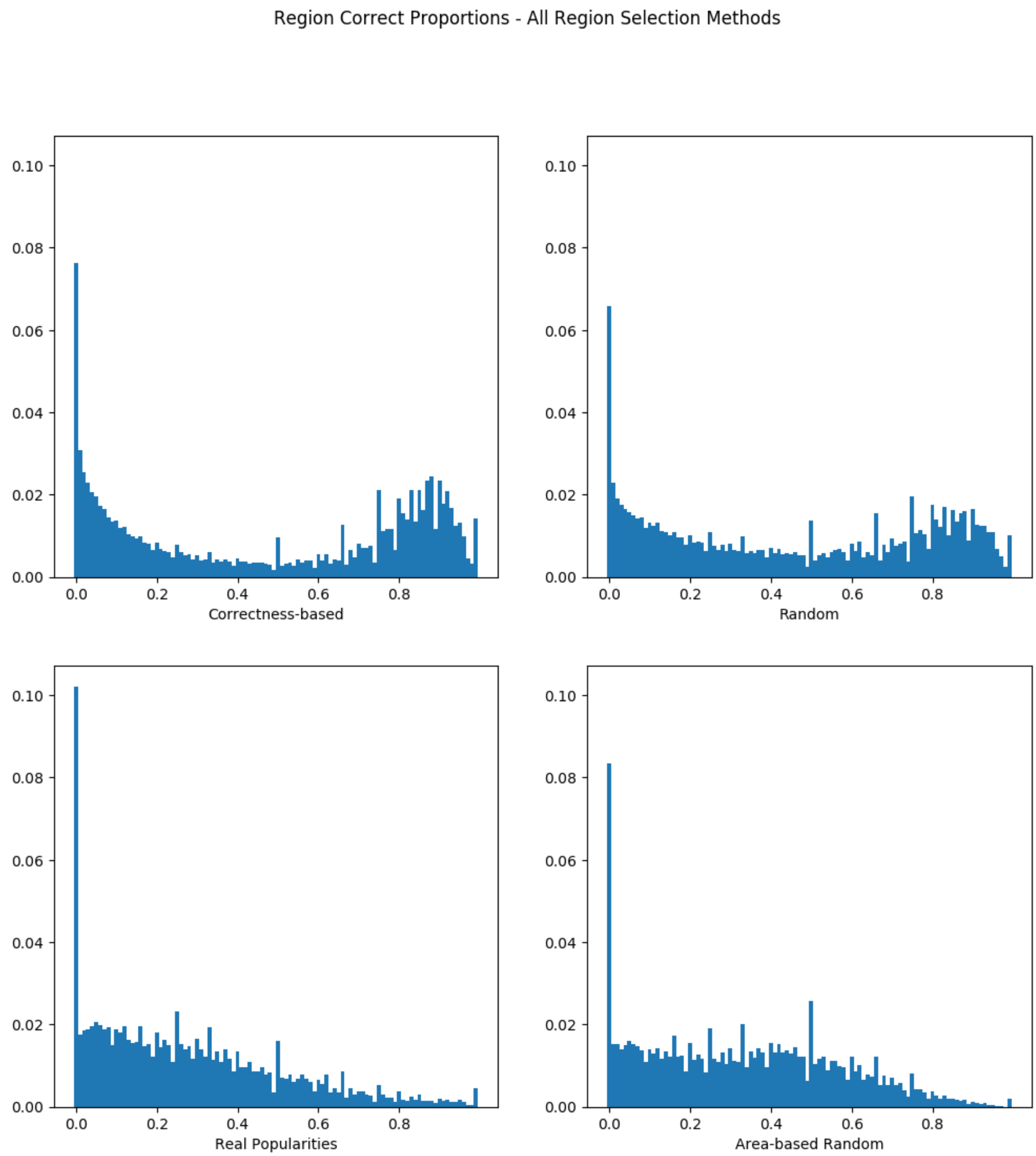


Figure 5.12: The distribution of region correctness proportions with each Region Selection method.

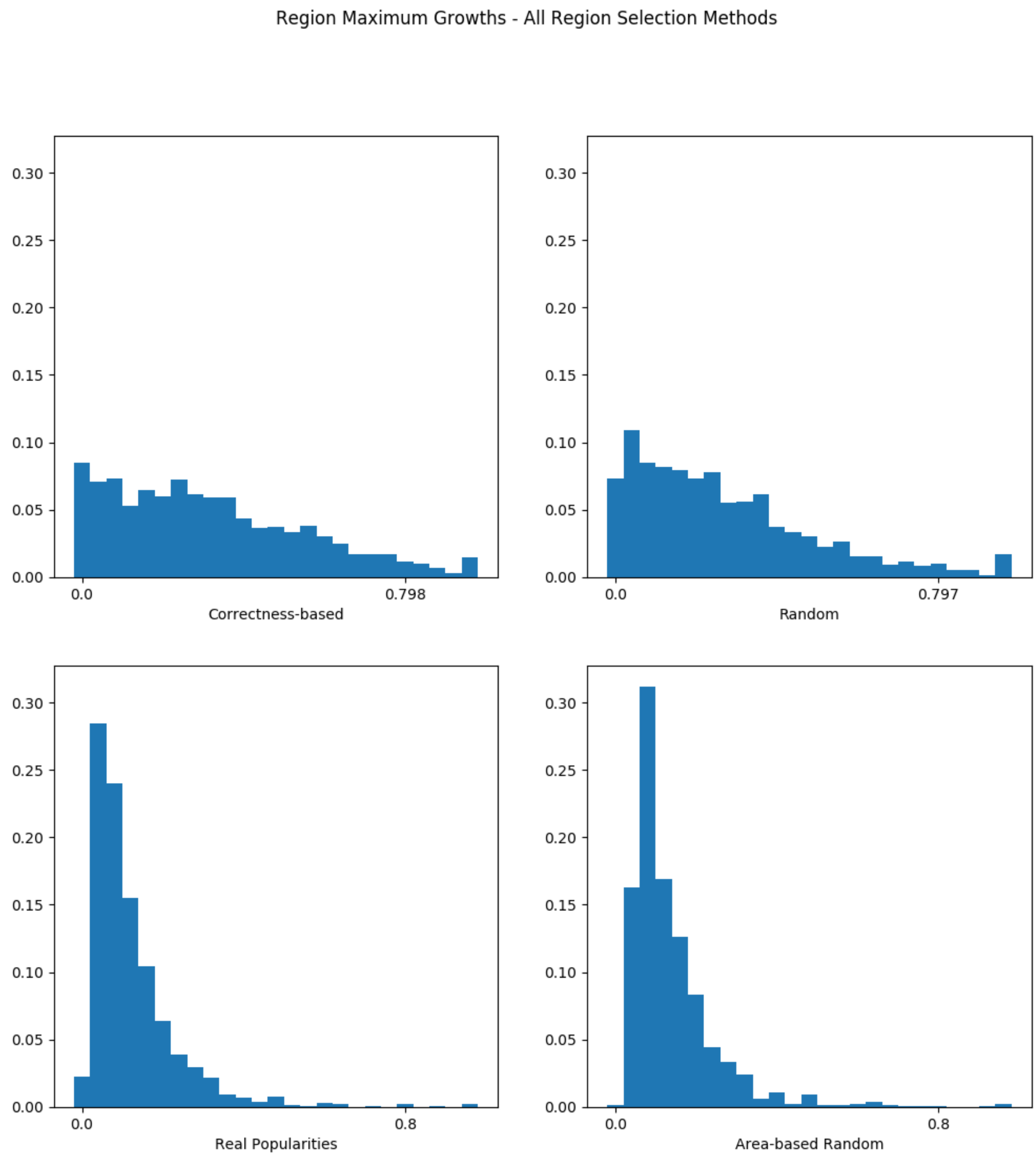


Figure 5.13: The distribution of region maximum growth amounts with each Region Selection method.

Region Selection Method	Cosine Similarity
Correctness-based	0.866
Area-based Random	0.770
Random	0.855
Real Popularities	0.826

Table 5.12: Average Cosine Similarity between each Region selection method and experiment data for the Region Correct Proportions metric.

Region Selection Method	Cosine Similarity
Correctness-based	0.812
Area-based Random	0.348
Random	0.717
Real Popularities	0.323

Table 5.13: Average Cosine Similarity between each Region selection method and experiment data for the Region Maximum Growths metric.

methods in Figure 5.13. The numerical comparison of the Maximum Growth metric is shown in Table 5.13.

5.4 Discussion

In this chapter we have developed an agent-based model to simulate the /r/Place experiment with the goal of gaining a better understanding of the causes behind a number of outcomes of the experiment. We began by building our model and tuning several parameters to reach a high level of accuracy when compared with data from the original experiment. Our attention then focused on two areas: the role that coordination, either internal or external, played in the development of Regions, and the mechanism impacting which Regions to which users chose to contribute.

First, we studied the effect that varying the strength of coordination had on Region development. In our initial experiments we disabled coordinators, one type at a time, and found there was no noticeable effect on the simulation. This was likely due to our definitions of coordination being slightly restrictive so that the true level of coordination was underrepresented in the model.

When we increased the prevalence of coordination we observed some interesting changes in the results. First, creating greater numbers of Coordinator Agents caused a significant increase in the proportion of Regions that were made entirely correct. This is not unexpected; the Regions that become coordinated receive a rapid influx of correct pixels that are very likely to cover any incorrect pixels so they will quickly become highly correct.

Similar results are observed when we strengthen External Coordination by causing an order-of-magnitude increase in the number of pixels placed during a period of External Coordination in a Region. In the original experiment, we believe such a period corresponds to an outside force compelling non-participating Reddit users to contribute to a specific Region. More pixels placed from External Coordination corresponds to a more visible or popular external force. In our simulations when we increase the strength of External Coordination we see an increase in the proportion of Region correctness samples that are approximately 90-99% correct but not 100% correct. Increasing the strength of Internal Coordination, on the other hand, causes an increase in the fraction of Region correctness samples that are 100% correct.

What this indicates is that Internal Coordinator agents tend to cause much more rapid Region development than External Coordinator agents since Regions affect by Internal Coordination spend much less time *almost* complete and more time entirely complete. The parallel for this effect in the /r/Place experiment is that Regions which inspire many users to work on building just that Region will tend to grow more quickly than Regions gaining large amounts of support from outside the experiment.

It seems that Coordination, in either of the forms modeled in this chapter, does not appear to have a profound impact on the experiment as a whole. However, it can greatly enhance the development of the few Regions fortunate enough to benefit from it. This result is quite interesting as it seems to, perhaps, suggest some form of collective or group intelligence responsible for the creation of so many clearly designed Regions. The mechanism behind how this happened remains unclear.

We also studied the manner in which users may have chosen where to place their pixels. Four Region Selection methods were considered and we found that while our default method of choosing Regions (which favoured Regions that were partially correct over those entirely correct or incorrect) performed better than other metrics the difference between it and random chance was quite slim.

The fact that our Region Selection method performed fairly well suggests that it is at least partially correct and users had some tendency to contribute more to Regions that were partially correct. It may be that the large number of variables affecting the human decision-making process made the method nearly indistinguishable from random chance. We did certainly observe, however, that selecting Regions based on Area seems to perform quite poorly suggesting that the area of a Region does not play a large role in determining how many users choose to contribute to the Region. We also saw that using the Region selection chances based on real data performed poorly. This may be due to the various simulation parameters that we changed, which likely varied the simulation state enough to render the real data useless.

Chapter 6

Conclusions

In this thesis we have introduced the Reddit Place Experiment, an exciting and previously unstudied dataset of online collaborative human interaction in a novel setting. The experiment consisted of a combination of human users and bots acting collaboratively, randomly, and antagonistically to place pixels upon a canvas. These users developed either through communication, or spontaneously, over 1000 Regions each representing some particular interest.

In Chapter 4 we considered /r/Place as a peer production platform and compared it with a number of previous results from research into Wikipedia. Across a number of metrics, /r/Place contributors tend to behave similarly to Wikipedians. Specifically, we saw that membership turnover, the process of existing users departing and new users arriving, tends to slightly increase productivity. As well, users who contributed to a smaller number of projects tended to provide more helpful contributions than users who spread their efforts. Interestingly, we observed the largest departure from previous results when comparing aspects of the two platforms that are more reliant on the platform domain itself: User behaviour appears similar but Regions in the Place Experiment grew differently than Wikipedia pages.

Constructing an agent-based model of /r/Place in Chapter 5 gave us insight into how coordination affected the outcome of the experiment. Our ABM was designed to investigate two different types of coordination: spontaneous coordination, in which users did not engage in any sort of explicit planning with other users and external coordination, in which users designed very specific Regions outside of /r/Place and built them together. Contrary to our expectations, we found limited evidence that either type of coordination necessarily had a significant effect on the outcome of Regions that were not themselves being coordinated. This pointed at either an incomplete definition of coordination, or the manifestation of some unidentified form of collective intelligence.

6.1 Future Work

This is the first analysis of what we hope will be many attempts to glean information about crowd behaviour from the Place Experiment. As such, there is a large amount of work that could be done in the future.

First, we believe much stronger results might be obtained with the use of an expanded set of labels. The Place Atlas [41] that we use throughout this thesis provides a labeled set of Regions that are present at the conclusion of the experiment. Being able to track the development of these Regions throughout the entirety of the experiment would be invaluable to future analysis. Initial attempts to automatically track Regions across time using clustering methods and some domain knowledge have met some limited success.

Additional data could also be acquired via the Reddit API. By searching for posts relating to coordination in /r/Place, additional information about each coordinated Region could be discerned, such as an estimate of how many people saw each coordination post, or when it was posted in relation to the Region's growth. Using this data could cast additional light on the circumstances leading up to external coordination in Regions.

While we have observed several similarities between contributors to /r/Place and Wikipedia editors in Chapter 4, we have not made any comparison with other domains. Analyzing the behaviour of users of alternative peer production platforms, such as contributors to open source software projects, is an important next step in understanding which trends generalize across domains and which do not. This could aid in the design of future peer production platforms and crowdsourcing platforms or tasks.

We hypothesized that a large reason we saw little effect from coordinating users in Chapter 5 was due to the weak definition of coordination that we used. A broader definition of coordination might allow for future analysis to more accurately identify the effects of coordination throughout the experiment, or more thoroughly suggest that coordination had little impact.

Finally, in our ABM we have used a relatively simplistic model of human behaviour. A more advanced model might be created where agents take into account factors not used in our model such as where they have previously placed pixels, or the types of Regions to which they contribute.

References

[1] Robert Axelrod and William Donald Hamilton. The evolution of cooperation. *science*, 211(4489):1390–1396, 1981.

[2] Robert L. Axtell. Coordination in transient social networks: An agent-based computational model of the timing of retirement robert l. axtell and joshua m. epstein. *Generative social science: Studies in agent-based computational modeling*, 146, 2006.

[3] Martina Balestra, Lior Zalmanson, Coye Cheshire, Ofer Arazy, and Oded Nov. It was fun, but did it last? the dynamic interplay between fun motives and contributors activity in peer-production. *Human-Computer Interaction*, 1:1, 2017.

[4] Ralph Beckers, O.E. Holland, and Jean-Louis Deneubourg. From local actions to global tasks: Stigmergy and collective robotics. 1994.

[5] Yochai Benkler. Freedom in the commons: Towards a political economy of information. *Duke LJ*, 52:1245, 2002.

[6] Yochai Benkler. Peer production and cooperation. *Handbook on the Economics of the Internet*, 91, 2016.

[7] Thomas Berger and Pepijn Schreinemachers. Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society*, 11(2), 2006.

[8] Kim Bloomquist. Tax compliance as an evolutionary coordination game: an agent-based approach. *Public Finance Review*, 39(1):25–49, 2011.

[9] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287, 2002.

[10] T.G. Bransden and David G. Green. Getting along with your neighbours-emergent co-operation in networks of adaptive agents. In *Workshop on Intelligent and Evolutionary Systems (IES2005),. Future University-Hakodate, Japan*, 2005.

[11] Brian S. Butler. Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information systems research*, 12(4):346–362, 2001.

[12] Filippo Castiglione. Agent based modeling. *Scholarpedia*, 1(10):1562, 2006.

[13] Garrett M. Dancik, Douglas E. Jones, and Karin S. Dorman. Parameter estimation and sensitivity analysis in an agent-based model of leishmania major infection. *Journal of Theoretical Biology*, 262(3):398–412, 2010.

[14] Drunken_Economist. Place datasets (april fools 2017), 2017.

[15] David C. Earnest. Coordination in large numbers: An agent-based model of international negotiations. *International Studies Quarterly*, 52(2):363–382, 2008.

[16] Martin Ebner and Andreas Holzinger. Lurking: An underestimated human-computer phenomenon. *IEEE MultiMedia*, (4):70–75, 2005.

[17] Tom P. Evans and Hugh Kelley. Multi-scale analysis of a household level agent-based model of landcover change. *Journal of environmental management*, 72(1-2):57–72, 2004.

[18] W. Wayne Fu and Clarice C. Sim. Aggregate bandwagon effect on online videos' viewership: Value uncertainty, popularity cues, and heuristics. *Journal of the American Society for Information Science and Technology*, 62(12):2382–2395, 2011.

[19] John Geanakoplos, Robert Axtell, J Doyne Farmer, Peter Howitt, Benjamin Conlee, Jonathan Goldstein, Matthew Hendrey, Nathan M. Palmer, and Chun-Yi Yang. Getting at systemic risk via an agent-based model of the housing market. *American Economic Review*, 102(3):53–58, 2012.

[20] Simon Goss, Serge Aron, Jean-Louis Deneubourg, and Jacques Marie Pasteels. Self-organized shortcuts in the argentine ant. *Naturwissenschaften*, 76(12):579–581, 1989.

[21] Mary L. Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. The crowd is a collaborative network. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pages 134–147. ACM, 2016.

[22] Neha Gupta, David Martin, Benjamin V. Hanrahan, and Jacki O'Neill. Turk-life in india. In *Proceedings of the 18th International Conference on Supporting Group Work*, pages 1–11. ACM, 2014.

[23] Dirk Helbing. Agent-based modeling. In *Social self-organization*, pages 25–70. Springer, 2012.

[24] Eunice Jun, Gary Hsieh, and Katharina Reinecke. Types of motivation affect study selection, attention, and dropouts in online experiments. *ACM Human-Computer Interaction 1, CSCW*, 2017.

[25] Raghav Pavan Karumur, Tien T. Nguyen, and Joseph A. Konstan. Early activity diversity: Assessing newcomer retention from first-session activity. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 595–608. ACM, 2016.

[26] Aniket Kittur, Bryant Lee, and Robert E. Kraut. Coordination in collective intelligence: the role of team structure and task interdependence. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1495–1504. ACM, 2009.

[27] Aniket Kittur, Bryan Pendleton, and Robert E. Kraut. Herding the cats: the influence of groups in coordinating peer production. In *Proceedings of the 5th international Symposium on Wikis and Open Collaboration*, page 7. ACM, 2009.

[28] Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. He says, she says: conflict and coordination in wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 453–462. ACM, 2007.

[29] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan. Mason: A multiagent simulation environment. *Simulation*, 81(7):517–527, 2005.

[30] Meng Ma and Ritu Agarwal. Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information systems research*, 18(1):42–67, 2007.

[31] Jennifer Marlow, Laura Dabbish, and Jim Herbsleb. Impression formation in online peer production: activity traces and personal profiles in github. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 117–128. ACM, 2013.

[32] Brian James McInnis, Elizabeth Lindley Murnane, Dmitry Epstein, Dan Cosley, and Gilly Leshed. One and done: Factors affecting one-time contributors to ad-hoc online communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 609–623. ACM, 2016.

[33] Audris Mockus, Roy T. Fielding, and James D. Herbsleb. Two case studies of open source software development: Apache and mozilla. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 11(3):309–346, 2002.

[34] David S. Moore, William Notz, and Michael A. Fligner. *The basic practice of statistics*. WH Freeman, 2013.

[35] Scott Moss. Alternative approaches to the empirical validation of agent-based models. *Journal of Artificial Societies and Social Simulation*, 11(1):5, 2008.

[36] Richard Nadeau, Edouard Cloutier, and J.H. Guay. New evidence about the existence of a bandwagon effect in the opinion formation process. *International Political Science Review*, 14(2):203–213, 1993.

[37] Katherine Panciera, Aaron Halfaker, and Loren Terveen. Wikipedians are born, not made: a study of power editors on wikipedia. In *Proceedings of the ACM 2009 international conference on Supporting group work*, pages 51–60. ACM, 2009.

[38] Jenny Preece, Blair Nonnecke, and Dorine Andrews. The top five reasons for lurking: improving community experiences for everyone. *Computers in human behavior*, 20(2):201–223, 2004.

[39] Reid Priedhorsky, Jilin Chen, Shyong Tony K Lam, Katherine Panciera, Loren Terveen, and John Riedl. Creating, destroying, and restoring value in wikipedia. In *Proceedings of the 2007 international ACM conference on Supporting group work*, pages 259–268. ACM, 2007.

[40] Sam Ransbotham and Gerald C. Kane. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in wikipedia. *Mis Quarterly*, pages 613–627, 2011.

[41] Roland Rytz. The place atlas, 2017.

[42] Amit Singhal. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4):35–43, 2001.

[43] Peter Stone, Gal A Kaminka, Sarit Kraus, and Jeffrey S Rosenschein. Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *AAAI*, 2010.

[44] Jan C. Thiele, Winfried Kurth, and Volker Grimm. Facilitating parameter estimation and sensitivity analysis of agent-based models: A cookbook using netlogo and r. *Journal of Artificial Societies and Social Simulation*, 17(3):11, 2014.

[45] Bogdan Vasilescu, Daryl Posnett, Baishakhi Ray, Mark G.J. van den Brand, Alexander Serebrenik, Premkumar Devanbu, and Vladimir Filkov. Gender and tenure diversity in github teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3789–3798. ACM, 2015.

[46] Bogdan Vasilescu, Alexander Serebrenik, Mathieu Goeminne, and Tom Mens. On the variation and specialisation of workloada case study of the gnome ecosystem community. *Empirical Software Engineering*, 19(4):955–1008, 2014.

[47] Michael Wunder, Siddharth Suri, and Duncan J. Watts. Empirical agent based models of cooperation in public goods games. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 891–908. ACM, 2013.

[48] Ming Yin, Mary L. Gray, Siddharth Suri, and Jennifer Wortman Vaughan. The communication network within the crowd. In *Proceedings of the 25th International Conference on World Wide Web*, pages 1293–1303. International World Wide Web Conferences Steering Committee, 2016.

[49] Ming Yin and Yu-An Sun. Human behavior models for virtual agents in repeated decision making under uncertainty. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 581–589. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[50] Bowen Yu, Yuqing Ren, Loren Terveen, and Haiyi Zhu. Predicting member productivity and withdrawal from pre-joining attachments in online production groups. In *2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017*. Association for Computing Machinery, 2017.

[51] Bowen Yu, Xinyi Wang, Allen Yilun Lin, Yuqing Ren, Loren Terveen, and Haiyi Zhu. Out with the old, in with the new? unpacking member turnover in online production groups. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW), 2017.

[52] Haifeng Zhang and Yevgeniy Vorobeychik. Empirically grounded agent-based models of innovation diffusion: a critical review. *Artificial Intelligence Review*, pages 1–35, 2017.

[53] Haifeng Zhang, Yevgeniy Vorobeychik, Joshua Letchford, and Kiran Lakkaraju. Data-driven agent-based modeling, with application to rooftop solar adoption. *Autonomous Agents and Multi-Agent Systems*, 30(6):1023–1049, 2016.