석 사 학 위 논 문

Master Thesis

시민참여형 시스템에서
개인화가 참여에 미치는 영향

Effects of Personalization on Participation in
Crowd-civic Systems

2018

Grau, Paul

한 국 과 학 기　　원

Korea Advanced Institute of Science and Technology

Effects of Personalization on Participation in Crowd-civic Systems

Grau, Paul

Advisor: Kim, Juho

A thesis submitted to the faculty of
Korea Advanced Institute of Science and Technology
in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science.

Daejeon, Korea

May 31, 2018

Approved by

_____

Kim, Juho, Ph.D.

Assistant Professor

The study was conducted in accordance with Code of Research Ethics.

I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my thesis contains honest conclusions based on my own careful research under the guidance of my advisor.

Bibliographic Information

초록   군중 시민 시스템(Crowd-civic system)에서는 시민들이 집단을 형성하며 사회적 이슈를 발견하거나 공식 정책을 개혁하는 것과 같은 공동 목표를 향해 일한다. 불행하게도 많은 실제 시스템이 게임화 및 설득 기법과 같은 다양한 접근 방식에도 불구하고 많은 수의 시민들이 자발적으로 참여하도록 동기를 부여하지 못 한다. 이 논문에서는 자기 결정 이론(Self-determination theory, SDT)에 의해 주장 된 동기 부여를 지원하기 위해 고안된 개인화 된 메시지의 영향을 조사한다. 커뮤니티 문제를 수집하기 위한 개인화 된 동기 부여 지원 메시지를 표시하는 군중 시민 플랫폼을 설계하고 두가지 실험을 수행했다: Amazon Mechanical Turk에서 150 명의 참가자와의 쌍 비교 실험과 120 명의 대학 구성원과의 필드 스터디. 쌍 비교 결과는 군중 시민 체계에서 SDT의 관점을 적용 할 수 있음을 나타냈다. 실제 시스템에서 적용하는 동안 일반 동기를 방해하지 않고 참가자를 모집하는 등 몇 가지 과제가 제기되었지만 수집 된 데이터는 비슷한 유망한 추세를 나타냈다.

핵심 낱말   시민기술, 크라우드소싱, 동기부여, 참여, 개인화

Abstract   In crowd-civic systems, citizens form groups and work towards shared goals, such as discovering social issues or reforming official policies. Unfortunately, many real-world systems have been unsuccessful in continually motivating large numbers of citizens to participate voluntarily, despite various approaches such as gamification and persuasion techniques. In this thesis, I examine the influence of personalized messages designed to support motivation as asserted by the Self-Determination Theory (SDT). I designed a crowd-civic platform for collecting community issues with personalized motivation-supportive messages and conducted two studies: a pair-comparison experiment with 150 participants on Amazon's Mechanical Turk and a live deployment study with 120 university members. Results of the pair-comparison study indicate applicability of SDT's perspective in crowd-civic systems. While applying it in the live system surfaced several challenges, including recruiting participants without interfering with general motivations, the collected data exhibited similar promising trends.

Keywords   Civic technology, crowdsourcing, motivation, participation, personalization

Table of Contents

Abbreviations

SDT   Self-Determination Theory
MTurk   Amazon Mechanical Turk
GCOS   General Causality Orientations Scale
MVS   Motivation to Volunteer Scale

1

Introduction

Interest in civic technology has been growing quickly in the past decade. Researchers have been seeking novel technical solutions to fundamental problems of communities: how to "govern, organize, serve, and identify matters of concern" [4]. In crowd-civic systems, large numbers of citizens form groups in processes that are related to crowdsourcing [32]: they join online communities and contribute small pieces towards common goals, such as the collection of local knowledge, discovery of social issues, or even reforming official policies. Users of crowd-civic systems work together to achieve tangible outcomes, and they often do so voluntarily.

Encouraging people to join an activity and keeping them motivated has always been a major challenge for crowdsourcing system designers [43]. Even in paid micro-task crowdsourcing, in which direct influence of external rewards is expected, research has shown that a combination of factors determines the workers' decision to opt in [38]. In crowd-civic systems, citizens are typically asked to contribute voluntarily for a variety of reasons, including the greater good (e.g. as a civic duty) and other indirect benefits (e.g. social or epistemic aspects) [2], in absence of monetary rewards. Thus, due to diverse motivations of citizens [2], the issues of self-selection [1] and the long-tail of participation [53], often a particular, small, and homogeneous group of citizens ends up contributing most of the work. This is less than desirable for fulfilling the democratic ideal of representativeness [1]. For technology to advance social progress, all voices need to be heard. However, there is a large and underrepresented group of people that participate rarely or never in these kinds of systems, which has lead researchers to call for more inclusive, differentiated design — for moving away from the currently prevalent "one size fits all designs of civic technology" [2].

In the field of work motivation, theories often differentiate between *intrinsic* and *extrinsic* motivation based on the origin of the motivation. Intrinsic motivation is rooted internally and leads to direct satisfaction just through the performance of an activity itself, whereas satisfaction from extrinsic motivation is rooted in separable consequences of the activity, such as monetary rewards [12]. Despite the positive effect of intrinsic motivation on performance [12], in reality, many tasks are not composed to produce intrinsic motivation. The Self-Determination Theory (SDT) [44] is a well-established theory of motivation which is empirically validated in different domains like education, health care, and work.

It distinguishes between intrinsic motivation and spectra of extrinsic motivation in terms of involved behavior regulation processes. Besides the task itself, SDT also considers the individuals who perform an activity and how they internalize its reasons and goals. It provides a unique framework for studying the influence of user and environment factors, as well as for investigating their interaction on motivation and, consequently, on desired outcomes.

Previous research has shown how users with varying personalities, as described by standard tests like the Big Five personality traits, have a preference for different motivational affordances [19]. This is an example for "personality-targeted design" [39], a framework for theory-driven personalized user interfaces. Addressing individual differences can be effective to increase participation, not only in total but also from more diverse people, and as a result, boost representativeness [20]. To my knowledge, this idea has not been investigated empirically in the domain of crowd-civic systems. Thus, in this work, I address the research question: in what way can motivation theory inform the design of personalized user interfaces of crowd-civic systems to support different motivations?

My work is centered on the thesis that **personalized motivation-supportive design has positive effects on participation in a crowd-civic system**. I conducted two studies to gather empirical evidence for this claim. Results of a controlled pair-comparison study indicate the general applicability of Self-Determination Theory (SDT) in crowd-civic systems and the possibility of personalization. Data collected in a subsequent live deployment using an application designed based on the initial findings exhibited similar promising trends, but several challenges arose regarding personalization. By studying this concept in two settings, one focusing on control and one on ecological validity, I want to provide a more balanced view on the possible advantages and disadvantages.

This document is organized as follows:

In Related Work, I explain the theoretical foundations and previous research that this study is built upon. Research Questions gives an overview over the two studies that I conducted. Study 1: Online Survey with Crowd Workers introduces the method and results of the first study. In Design of a Research Probe, I present the design of the case study application that is then used in Study 2: Field Study. This section is followed by an in-depth Discussion of all results. Finally, I summarize the findings and point out possible future research directions in Conclusion and Future Work.

2

Related Work

In this chapter, previous research that is related to this thesis is presented. First, I introduce the application domain of civic technology and its challenges. Then, I talk about some general strategies for increasing user engagement in computer-supported cooperative systems. In particular, I focus on the possibility of using personalization for this purpose. Finally, I introduce the Self-Determination Theory (SDT) of motivation which the design of this study is heavily based on.

2.1 Civic Technology

2.1.1 Overview

Civics, broadly speaking, encompasses everything related to the rights and duties of citizens. Technology is increasingly used to improve and redefine civics [4]. For example, one aspect that has now become mainstream for many democratic governments is trying to improve transparency by enabling citizens to better learn and understand what the government is doing. Under the Open Government Partnership, 75 countries have officially committed to making their governments more open and accountable. Other kinds of civic technology include public deliberation, voting, local sharing, and information crowdsourcing. This encompasses systems that are uni-directional (i.e. government to or from citizen), bi-directional (i.e. government to citizen and vice-versa), or multi-directional (i.e. communities of multiple stakeholders). Figure 2.1 shows clusters of recent trends in this field.

There have been a multitude of studies investigating the use of technology in the civic space from various angles, such as social science, political science, design, and computer science.

The prevailing assumption among researchers is that "new technologies will, in some way, shape, or form, be desirable and advance social progress" [4], but at the same time, technology's limitations and flaws are also questioned. Technologies such as smartphones or IoT devices could increasingly empower citizens to the end of collecting their own data and connecting themselves. This could enable a future in which governments shift from a provider role to a coordinator role. Moreover, information might
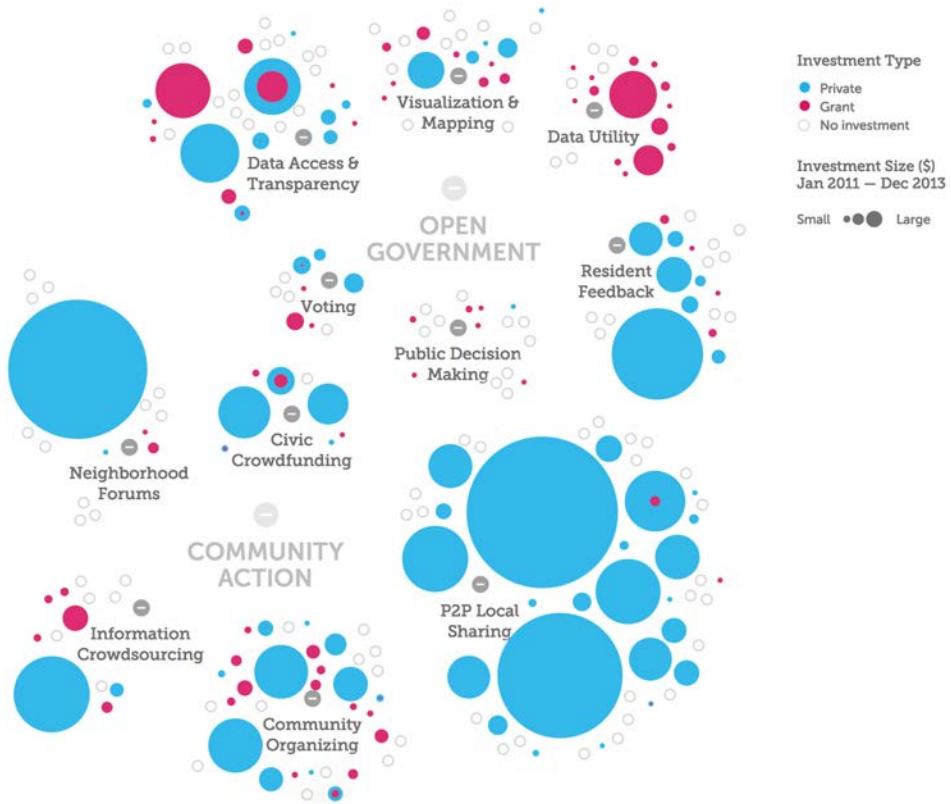
Figure 2.1: Trends in Civic Tech (© Knight Foundation, 2014 [42]). The visualization shows topics by investment size, arranged in two main clusters: open government and community action.

be increasingly "flattened", i.e. freed from the traditional hierarchical organization into a form more easily accessible, dubbed the "Google effect" by the authors. Boehner and DiSalvo [4] propose that these technological advancements can ultimately also change how governments organize themselves.

Many researchers hope that advances in technology increasingly lower the barrier of citizen participation and enable effective communication within communities of any scale. Regarding empowerment of citizens, studies have shown that technology usage has positive effects on visibility, accountability, and accessibility [11]. However, it is also argued that "technology alone cannot equitably empower citizens" [11], because it cannot replace social ties and political power. Technology is not a silver bullet.

2.1.2 Crowdsourcing

Crowdsourcing is a model for obtaining results through a group of online users [43]. Typically, crowd members work together by contributing small pieces that are then aggregated to achieve a shared goal. Crowdsourcing can be seen as a distributed and often asynchronous and depersonalized elicitation of responses from a large number of users. For over a decade, researchers have been thoroughly studying workflows to improve the quality of responses and the efficiency, effectiveness, and also worker experience of these tasks.

Many crowdsourcing systems rely on paying workers to complete requested tasks. Examples include are micro work tasks such as data labeling and providing specific information on platforms like Amazon Mechanical Turk (MTurk). However, in recent years, researchers have been using the crowdsourcing framework in novel domains such as education (learnersourcing [23]) and civic engagement (citizensourcing [50] or crowd-civic systems). In these systems, humans often engage voluntarily. Here, it is crucial to design proper incentives.

Many learnersourcing applications not only help users learn better, but their contributions also help future learners [23]. Some notable examples for citizensourcing include Ushahidi [40], BudgetMap [24], and Lab in the Wild [labinthewild.org]. Ushahidi employs crowdsourcing to map incidents of violence during a crisis. BudgetMap asks the crowd to label government data to make it more accessible. In Lab in the Wild, participants provide personal data through engaging in online experiments that are intrinsically interesting. What these have in common is that they try to motivate users in two ways: doing a task that is beneficial for themselves, but also a meaningful contribution to a greater cause.

2.1.3 Crowd-civic Systems

Crowd-civic systems are an emerging category of civic technology applications where citizens form groups working towards shared goals, such as the collection of local knowledge, discovery of social issues, or reforming official policies. It is the application of crowdsourcing to civic contexts that comes with a number of unique challenges [32]. A major issue of crowd-civic systems is attracting and retaining the "right" participants: depending on the application, different characteristics of participants might be desirable, such as expertise, diversity of opinions, and representativeness. The plain number of volunteers is also a critical issue. After an initial "hype", maintaining a critical mass for long-term engagement has been difficult for many platforms [53].

Aitamurto and Landemore [1] conducted a study in Finland employing elements of crowdsourcing for deliberation on lawmaking. In an open call, citizens were invited to join an online discussion forum to debate details of a proposed law change regarding off-road traffic. The researchers saw high user engagement that exhibited qualities of democratic deliberation as defined by previous research. However, the authors also note the unsolved problem of limited representativeness, owing mostly to the nature of self-selection and required level of technology skills.

Crowd-civic systems suffer from self-selection and non-diverse participants who aren't easily motivated to participate equally and for an extended period of time. This has lead numerous researchers to call for more inclusive, differentiated design. Combining these ideas while using the self-determination theory of motivation, my work tries to contribute a step towards moving away from the currently prevalent "one size fits all designs" [2].

2.1.4 Motivation of the Crowd

Motivations to participate in any task are usually very diverse, even within sub-categories of crowd-sourcing such as paid or voluntary work.

Even in paid crowdsourcing platforms, workers are motivated not only by monetary rewards but by a combination of factors, including the task's interestingness, expected workload, and profitability [38]. Studies have shown that the amount of workers' participation is positively related to their autonomous motivation. Moreover, Naderi [38] proposed that "workers with a specific major general motivation type perform better when the task is motivated by a same nature of motivation."

Participation in crowd-civic systems is often voluntary. Designers of these systems have to clearly define and present the benefits to the users. Some people might be motivated by acting "for the greater good", for altruistically advancing societies. Usually, personal motivations are more complex than altruism and vary significantly. In another study by Aitamurto and Saldivar [2], they analyzed the dynamic nature of motivation factors. Most of the factors fluctuated over time, coinciding with the stage the crowdsourcing process was at. The authors argue that design for crowdsourced policymaking should support epistemic and interactive aspects to keep participants engaged, and more generally, that system designers should take different motivations into consideration.

In their work about voluntary participation in online experiments, Jun et al. [20] showed how motivation affects selection of participants. They suggest to diversify promotions to engage users with different motivations or from different demographics, which could mitigate self-selection biases and reduce dropout.

This thesis extends this line of work by applying motivation theory and the idea of diversification in a civic context.

2.2 Strategies for Increasing Voluntary User Engagement

2.2.1 Gamification

Gamification is the use of game design elements in non-game contexts [10]. Examples are the use of points, badges, levels or leader boards. One well-known example is Foursquare [foursquare.com], an online community that incentivizes users to "check-in" to and rate locations to build a global database of places of interest. Checking in to places is rewarded by points, a high position on the leader board among one's friends, and becoming the "mayor" of a place.

Gamification can be seen as a particular kind of motivational affordances [52] that has been a critical component of the success of many commercial applications. It can also be a useful motivator in citizen science [41] or education [28].

However, gamification has many limitations. Not every context lends itself to being "gamified", for example regarding serious issues such as crime or law. Also, gamification pushes users towards a certain, mostly extrinsic motivation — getting points, being ranked highly, getting noted for achievements — which can be unnecessary or even harmful when the task at hand might also be intrinsically motivated. Some even argue that the overreaching gamification of everything leads to the dulling of the mind [3]. Thus, gamification cannot be the answer to all motivation problems.

2.2.2 Message and Incentive Diversification

As a perhaps more subtle effort than gamification, several works have investigated different methods of diversifying messaging strategies in various domains of voluntary action, such as social activism, behavior change, and online discussion.

Botivist [46] used bots on Twitter to individually address social media users that might contribute action to specific activism causes. The authors evaluated different strategies of messaging people with different message framings, such as referring to solidarity, loss, or gain. Surprisingly, the most direct call-to-action was on average more effective than other strategies of persuasion. However, the authors didn't account for possible individual differences between participants.

Kocielnik and Hsieh [25] explored diversifying message-based triggers in a behavior change application. To counter users' annoyance and boredom when seeing a message repeatedly, they generated target-diverse and self-diverse message variants, referring to concepts that are cognitively close to either the action or recipient of a message. They found that self-diverse messages worked well as reminders and even motivators, but they note that in terms of motivation, all their messages followed a single, generally positive motivational strategy. In my work, I explore messages which support diverse kinds of motivations that exist in individual users.

Investigating designs that encourage "one-time" or novice contributors of online communities, McInnis et al. [33] found that the phrasing of prompts matters. Small differences in call-to-action messages, such as referring to more or less specific goals, can affect the quantity and quality of contributions. The authors also used personal factors such as self-efficacy to draw correlations with both longer comments and higher responsiveness. They argue that personal factors could be used to achieve more representative recruiting. In my work, I extend this idea to providing personalized experiences within applications not only for recruiting, but for extended engagement within an application.

Related to personal motivations, Hsieh and Kocielnik [18] investigated how different incentives may attract different people. They tested incentives corresponding to different human values, such as openness-to-change and risk-seeking, and saw correlations between participants' choice of reward and self-reported values. The authors argue that using diverse incentives can improve participation diversity. In my work, I investigate a similar idea, but based on self-determination. In addition to a paid crowdsourcing experiment, I also investigate a voluntary setting.

2.3 Personalization of User Interfaces

2.3.1 Overview

Personalization is a concept that is studied in many fields with varying perspectives and objectives, and as such it has hard to come up with an exact definition. Most generally speaking, personalization is *the use of personal characteristics known about a user to adapt their experience of interacting with a system* [22]. The use of personalization has been researched in-depth in domains such as ecommerce, education, online media, and social networks.

A more formal definition of personalization can be found in the work of Murthi and Sarkar [37]. They define it as a "*process of using a customer's information to deliver a targeted solution*" and outline several stages of that process:

1. **Learning** about the user, for example by

   - explicit elicitation, i.e. directly asking the user for input

   - implicit elicitation, i.e. analyzing interaction logs

   - inference, i.e. combining explicit and implicit data with heuristics or ontologies.

2. **Matching** between the user and content or strategies
3. **Evaluation** of the system

   - business goals (e.g. conversion rates, engagement)

   - user well-being (e.g. quality, usability, satisfaction)

User information that can be gathered and used for personalization includes demographic data (age, status, place of residence, income, etc.), personal preferences and interests, behavior directly observed from interactions (e.g. clicking on certain items), and other indirect measures. To devise a good matching method, one needs a user model (*who is the user?*), a task model (*what do they want to achieve?*), and possibly a content model (*what are characteristics of the content?*).

Personalization has been shown to have a range of positive effects. For example, perceived personalization significantly increases customers' intention to adopt by increasing cognitive trust and emotional trust [27]. Systems that seem familiar or understanding the user's intentions are generally seen as more enjoyable. In online commerce systems, personalization has been shown to significantly increase conversion rates [21, p. 115]. In an Information Systems perspective, personalization is about adapting content and presentation to be relevant to the processing goal [48]. Here, personalization has been shown to decrease the effort required in decision making.

Personalization can be incorporated into many parts of a system, but in this thesis I am focusing on user interfaces, and more specifically wording and presentation of call-to-action components.

2.3.2 Main Areas of Personalization
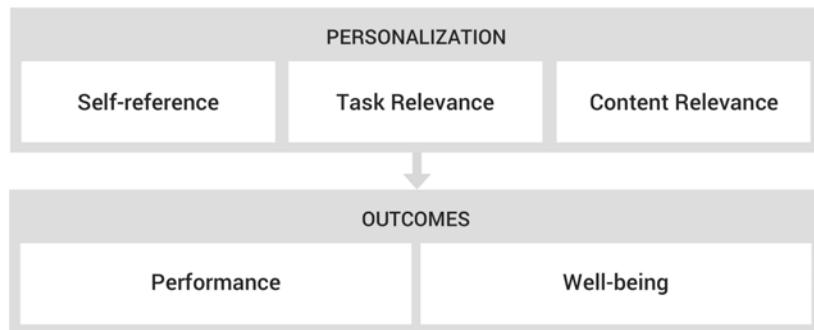
Figure 2.2: Three main areas of personalization and outcomes.

The three main directions of personalization that are generally referred to in literature are: self-reference, content relevance, and task relevance [48, 30].

- **Self-reference** means referring to a person's context whenever trying to convey any knowledge, task, or interaction. One domain where this kind of personalization has always been naturally employed is education: good teachers understand individual differences between their students and adapt their explanations and tasks according to their students' needs.

  In an experiment with elementary school children, contextualization, personalization, and choice all produced dramatic increases, not only in students' motivation but also in their depth of engagement in learning, the amount they learned in a fixed time period, and their perceived competence and levels of aspiration [5]. For example, an instruction that contains words like "you" or the reader's name is likely to be more interesting than a general text. When the wording of student materials is contextually consistent with the learner's interests, they learn better and enjoy their experience more [47]. The self-reference effect has been repeatedly shown as a major way of influence [48].

- **Content relevance** is perhaps what most people connect with personalization most easily. It is when Amazon recommends books that a user might like, Facebook predicts which stories someone might want to see, and Spotify suggests which music to listen to next. Content relevance is often predicted by recommender systems.

  Recommender systems are a field of study in Computer Science that is concerned with algorithms that predict a user's preference regarding an application's content or other users. Recommendations are perhaps the most widely deployed method of personalization. It can be found in countless examples across different domains, from ecommerce to content providers (news, music, videos), and social media.

- **Task relevance** means adapting the interaction context and navigational flow so that users can reach their goals more efficiently. Different users might use the same application or interface with different goals. Understanding their goals and offering shortcuts or clear indicators of what to do next can improve the user experience.

In this work, I am focusing on a kind of self-reference, or more specifically adapting the interface to *who the user is*. For that, I will introduce the notions of personality and personal motivation in subsequent sections.

2.3.3 Personality-targeted Design

Personality-targeted design, first introduced by Nov and Arazy [39], is a framework for theory-driven personalized user interfaces. In terms of the previously mentioned taxonomy, this relates to *self-reference*. For example, it was shown how a user's conscientiousness level can predict their perception of a "critical mass" [39]. The study UI contained an indicator of high or low participation from others, and the authors discovered a correlation between a user's conscientiousness level and the level of participation given one version of the interface or the other.

In a related study, Jia et al. [19] showed how people's personality, as measured with the Big Five test, affects their preference for different gamification affordances. For example, extroverted people thought that collecting points, advancing levels, and leaderboards are helpful and enjoyable, whereas people with higher levels of agreeableness were more likely to prefer challenges. Emotionally less stable participants disliked many of the affordances. In an older study, Moon [35] already suggested similar effects. In their study, participants completed a personality test to assign them to two groups: dominant and submissive. Then, they were asked to use an online shopping recommendation agent. Results showed that people were more likely to be persuaded by the recommendation when it was conveyed in a style matching their personality.

Gamification, or the use of game design elements in non-game contexts [10], has previously been employed to increase motivation, mostly for all users equally without taking into account individual differences.

These studies indicate new opportunities to use personalization in many parts of user interface design. Whereas so far personalization often meant predicting a user's interest in products or content, it can also be used to select strategies to make people engage more.

2.3.4 Limitations of Personalization

Despite its promised benefits, personalization has a number of possible limitations. In general, personalization effects are often hard to generalize, as they can vary based on individual difference characteristics or in different domains [47]. If personalization is done poorly, for example with irrelevant information,
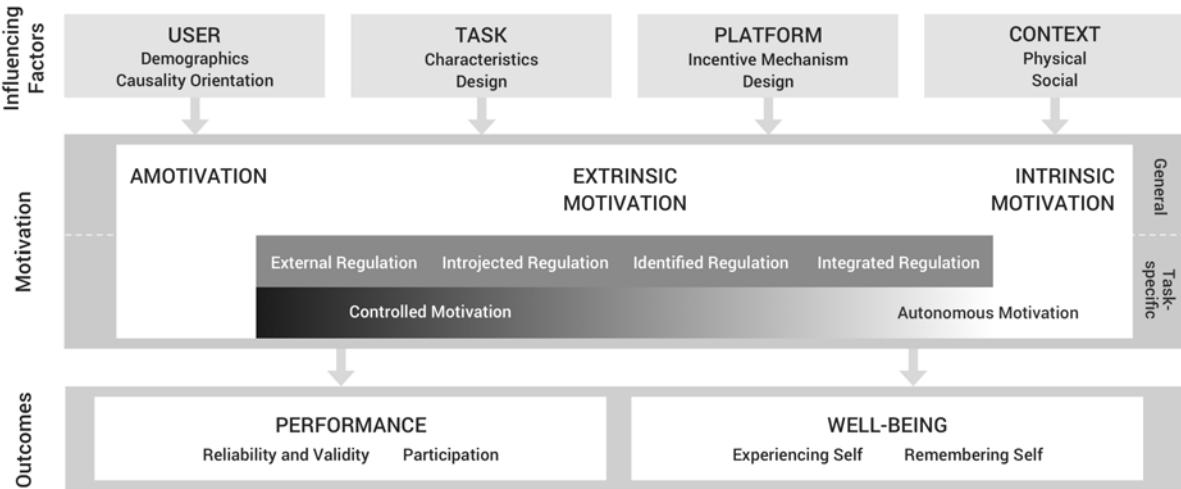
Figure 2.3: Taxonomy of motivation in crowd work, after Naderi [38, p. 8]. My work attempts to personalize the *platform* based on *user factors* to influence *participation* in a given *task*.

it can be more distracting than a non-personalized version [47]. More recently, in the context of filter bubbles and manipulation of public opinion, researchers have raised concerns that personalization may amplify bias [17]. Furthermore, as personalization typically needs a lot of personal data, there are several privacy concerns [6].

Given these problems, personalization is often a challenge of cost versus benefit. Cost includes both design, implementation, and maintenance costs for a company as well as effort and value of providing personal information for users. Benefits include increased conversion rates for a company and increased quality and usability for users.

2.4 Motivation-supportive Design

2.4.1 Self-determination Theory

SDT can offer numerous insights into designing with users' motivation in mind. SDT provides a broad, well-established, and empirically validated framework to study motivation. It has been developed by studying human behaviours and interactions in families, classrooms, and teams in different cultures. In this section, I provide a quick overview of SDT and how its findings relate to the context of crowd-civic systems.

SDT considers different types of motivations: intrinsic motivation, a spectrum of extrinsic motivation, and amotivation. Furthermore, it addresses the process of *internalization* in which "an individual acquires an attitude, belief, or behavioral regulation and progressively transforms it into a personal value, goal, or organization" [8]. Internalized types of extrinsic motivation share similar properties with intrinsic motivation. In combination, they create *autonomous motivation*, whereas other types of extrinsic motivation merges to create *controlled motivation*.

Within SDT's framework type and amount of someone's motivation do not only depend on the activity, but also to the person itself and the environment [44]. SDT claims that certain aspects of the environment can influence autonomous motivation. It also addresses three basic psychological needs and argues that they are universal for all humans: autonomy, competence, and relatedness (Table 2.2). Autonomy relates to "choosing to engage in behavior that is compatible with one's values, out of personal interest or expression of the self." However, autonomy alone is not sufficient. People require a sense of competence, or "feelings of effectiveness in one's work", to be effective. Finally, even with a complete sense of autonomy and competence, if a behavior happens in a vacuum without other people, it is usually not motivating enough. People need relatedness, i.e. they need "feeling connected with others, feeling interdependent with others, and feeling of belonging to a group or with other individuals.". When an environment supports the satisfaction of these needs, it is called *autonomy-supportive*. SDT claims that autonomy-supportive environments facilitate internalization of extrinsic motivation and positively influence performance and well-being of people [12].

Human behavior can be regulated by a variety of factors that can be mapped on a continuum of self-determination (Table 2.1). The processes involved in these regulations can be be thought in terms of an inverse relationship between self-determined and controlled motivation. However, these factors are environment- and task-specific — they can vary not just between individuals, but also between tasks of the same person. People's motivations are usually a combination of different regulations. In terms of individual orientations, SDT offers the concept of *general causality orientations*. They describe ways people generally and across many settings orient themselves to environments and regulate their behaviour (Table 2.3).

2.4.2 How to measure motivation

Measuring motivation directly is hard. In a recent study, participants of a crowd-civic systems were asked to self-report the personal relevance of a number of motivation factors [2]. In other studies, motivation has often been measured indirectly by other outcome metrics, such as level of interest, enjoyment, or taking on increasingly challenging tasks [5].

There are certain personality tests related to motivation. The General Causality-Orientation Scale (GCOS) [8] measures people's orientations on three scales: autonomy, control, and impersonal. A person who has a relatively high score on the autonomy scale values making their own choices and is less controlled by extrinsic rewards. Someone who is oriented towards control often acts because they think they have to through some form of external regulation or introjection of values. For such people, pressure in the form of deadlines or anticipated losses can play a bigger role. People who are high on the impersonal scale don't feel in charge of their own decisions, they like to follow precedents and often feel incompetent and not motivated.

Table 2.1: Behavior regulatory styles on the self-determination continuum, associated processes, and perceived locus of causality. Table reproduced from Ryan and Deci [44].

Amotivation	Extrinsic motivation				Intrinsic motivation
	External regulation	Introjection	Identification	Integration	
Perceived non-contingency, low perceived competence, nonrelevance, nonintentional-ity	Salience of extrinsic rewards or punishments, compliance, reactance	Ego involvement, focus on approval from self or others	Conscious valuing of activity, self-endorsement of goals	Hierarchical synthesis of goals, congruence	Interest, enjoyment, inherent satisfaction
Impersonal	External	Somewhat external	Somewhat internal	Internal	Internal
Lack of motivation	Controlled motivation	Moderately controlled motivation	Moderately autonomous motivation	Autonomous motivation	Inherently autonomous motivation

Table 2.2: Basic motivational needs according to SDT [44].

Need for autonomy	Need for competence	Need for relatedness
Choosing to engage in behavior that is compatible with one's values, out of personal interest or expression of the self.	Sense of proficiency and feelings of effectiveness in one's work	Feeling connected with others, feeling interdependent with others, and feeling of belonging to a group or with other individuals

Table 2.3: Individual general causality orientations according to SDT [8].

Impersonal orientation	Controlled orientation	Autonomy orientation
Believes that attaining desired outcomes is beyond his or her control and that achievement is largely a matter of luck or fate. Likely to be anxious and to feel very ineffective. They have no sense of being able to affect outcomes or cope with demands or changes. They tend to be amotivated and to want things to be as they always were.	Oriented toward being controlled by rewards, deadlines, structures, ego-involvements, and the directives of others. Likely to be dependent on rewards or other controls, and may be more attuned to what others demand than to what they want for themselves.	Oriented towards environments that stimulate intrinsic motivation, are optimally challenging, provide informational feedback, and allow choice. Tends to display greater self-initiation, seek activities that are interesting and challenging, and take greater responsibility for his or her own behavior.

The Motivation to Volunteer Scale (MVS) questionnaire [14] can be used to score a person on scales related to behavioral regulatory styles: Amotivation, External Regulation, Introjection, Identification, Integration, and Intrinsic Motivation.

The Crowdwork Motivation Scale (CWMS) [38] can be used to measure the general underlying motivation of crowd workers. It was developed and evaluated for paid crowd work, but could be adapted to other domains. However, in this thesis, I will be using MVS as its items are explicitly concerning voluntary participation while measuring similar constructs as CWMS.

2.4.3 Design based on Self-determination Theory

SDT has been used to inform design of interactive computer systems before. In his work, Deen [9] investigated how autonomy-supportive game design can improve motivation to learn. The author gives a wide range of examples of how findings from SDT can be translated to concrete design decisions. However, games make use of motivations which are quite distinct from voluntary participation in crowd-civic systems. The primary driver for games tends to be intrinsic motivation, e.g. enjoyment and inherent satisfaction, but voluntary participation is dependent on at least some extrinsic factors, e.g. importance of the issue and expected possible outcomes. In this thesis, I want to contribute to the discussion of supporting individual motivations in a voluntary participation context, crowd-civic systems.

Moreover, if one can support motivation by creating certain environments, and if motivation factors differ not just between tasks but also between individuals, this leads us to the idea of motivation-supportive, personalized design.

3

Research Questions

In Related Work, I have shown how motivation is a challenge in crowd-civic systems and how it is related to Self-Determination Theory (SDT) and personalization. I argued that crowd-civic systems require autonomous motivation, perhaps more than other systems. SDT offers rich insights into autonomous motivation that seem fit for personalization. In this chapter, I will define and explain the research questions that govern this work.

My overall research goal is finding out how findings from SDT can help designers of crowd-civic systems to create more engaging designs for diverse people. For this, I first want to gather some evidence that SDT-related concepts apply in this application domain. As a second step, I want to use this knowledge to build a personalization system with the goal of offering users affordances that support their individual motivations in order to increase users' engagement.

SDT has been studied in various domains of organizations, including in the context of organizational citizenship behaviors [12]. It has been argued that satisfaction of the three basic needs will enhance workers' motivations and in turn yield the outcomes deemed important in the respective field of work. In a volunteer work setting, perceived autonomy support may relate positively to the amount of volunteering. Interestingly, while studying factors for involvement in an important national referendum, it was found that intrinsic motivation itself is not enough to predict actual voting behavior. "When intrinsically interested in the issues, people became well informed, but only when they were motivated by the importance of the issues to themselves were they likely to actually exert the effort to go out and vote." [12, citing 26] Furthermore, numerous studies have shown the potentially detrimental effects of external rewards on autonomous motivation: payments decrease voluntary pro-social behavior [12]. This leads us to **RQ 1: What is the effect of external regulation on autonomously motivated online volunteer work in a crowd-civic system?**

For certain work environments (such as many other forms of crowdsourcing), it may be beneficial to focus on getting only the best participants and supporting their autonomous motivations. If diversity and representativeness is the goal, however, as in crowd-civic systems, that might neglect a large part of the population. It is also a matter of perspective of time: in the long term, supporting autonomous

motivations may be beneficial for everyone, but for contexts like a crowd-civic system that need to be effective immediately and sustainedly, we have to consider what we can do momentarily to support individuals to achieve the goal of increasing participation. This leads us to **RQ 2: Can individually targeted motivation support increase diverse engagement in crowd-civic systems?**

To address these two questions, two studies were envisioned. Study 1 was an online experiment conducted with paid crowd workers to gather initial evidence for the general applicability of the theory. Study 2 involved a live deployment study to evaluate the same idea in a more ecologically valid setting. In the second study, participants joined voluntarily with no expectation of being paid. By conducting these two studies, I aimed to provide a more balanced view on this topic.

Table 3.1: Overview over studies conducted for this thesis.

Study	Method	Data collection	Participants	Compensation	N
1	Online survey	Self reporting	Crowd workers	Paid	150
2	Field study	Usage data	University members	Voluntary	120
	— Post-survey	Self reporting	University members	Paid	38

4

Study 1: Online Survey with Crowd Workers

To gather initial evidence that using Self-Determination Theory (SDT) to inform messaging strategies of crowd-civic systems to support diverse motivations holds merit, I ran a study on Amazon Mechanical Turk (MTurk). I designed a user interface mockup for a fictional mobile application which features a prominent call-to-action to voluntarily participate in collecting ideas for a community. I implemented seven different message strategies based on SDT and collected rankings of individual preferences for each version through pairwise comparisons.

4.1 Hypotheses

The hypotheses for this study were:

H1: Different people prefer different motivation strategies, i.e. there is not a clear "winning strategy".

H2: Individual preferences correlate with personal differences in motivation or causality-orientation.

4.2 Design

This fictional application aims to collect a community's issues and ideas. I designed the main screen (Figure 4.1) which shows the feed of issues and a prominent call for participation.

I designed seven different versions of this call-to-action: a baseline that features only the basic instruction, and six versions that are hypothesized to support different kinds of motivations, as explained in detail in section 2.4. The versions are shown in Table 4.1. It is expected that the causality-orientated versions (**A**utonomy, **C**ontrol, **I**mpersonal) exhibit bigger differences overall, as A combines multiple aspects of the need-based versions, C is the only version with a strong extrinsic reward, and I also has very unique content. In addition, I included three versions each focusing on a particular basic need (nA, nR, nC) to investigate if there are any significant differences between them.

4.3 Method and Measures

I asked MTurk participants to (1) answer two questionnaires about general-causality orientation and motivation to volunteer, (2) rate their personal preferences between pairs of all interface versions, and (3) answer some questions about the design. The data collected also included basic demographic information (age, sex, education) and the frequency of volunteering, if any. The specific prompt for the pair comparison task was, "in which version would you personally be more likely to contribute a new idea?" Participants were asked to pairwise-compare all of the seven versions (i.e. they submitted $\binom{7}{2}$ + 1 = 22 A/B preferences covering all combinations and one repeated comparison as an attention check, seeing only one combination at a time).

For measuring relative preferences based on subjective attributes, pairwise comparison was used. The experiment was fully balanced: all participants made every possible comparison [7], and the order of the comparisons was balanced by the method of a Latin square.

Before seeing the design, participants were given the following scenario to aid their imagination in this hypothetical task, by pointing out the importance and potential impact:

> You joined an application that aims to gather ideas to improve conditions in your community, e.g. the members of your university, workplace, or neighborhood. We want as many people as possible to participate, because not everyone knows about the troubles of everyone else. The officials have agreed to monitor the community's ideas and follow-up with concrete steps once an issue has been identified.

Figure 4.1: Design sketch, baseline (BL) version. The screen includes a feed of recent ideas and a prominent call-to-action at the bottom, for which I designed 6 alternative versions.

After the comparison task, participants were presented with the "winner" according to their ratings and asked to give a one-sentence explanation of why they thought that this was the best version for them (or not).

For the test of general-causality orientation, the General Causality Orientations Scale (GCOS) questionnaire was used along with the recommended scale and ordering [8]. Through 12 questions with 3 possible responses each, participants rated the likelihood of responding to a scenario in a certain way on a scale of 1 to 7. These ratings are summed into three dimensions which represent orientations towards autonomy, relatedness, and competence.

Table 4.1: Study 1 design versions and messages for the prototype's call-to-action component following the rationales from Table 2.2. The messages are inserted into the screenshot depicted in Figure 4.1.
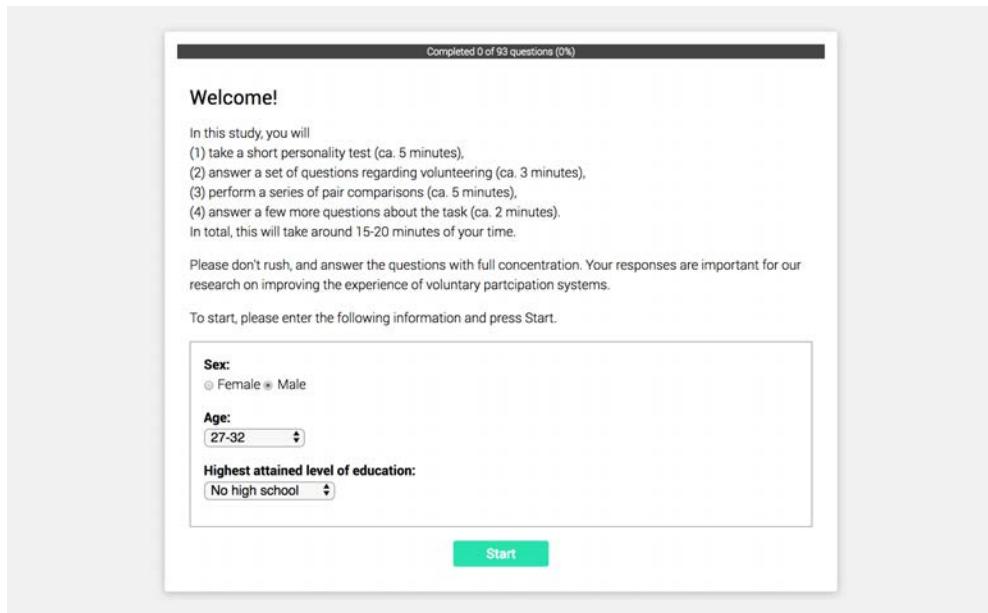
Need for autonomy (nA)	**Need for competence** (nC)	**Need for relatedness** (nR)
Sharing ideas matters! [Share your idea] Express your opinion, let everyone hear our voice!	Your opinion is needed. [Share your idea] You can have an impact on your surroundings.	Your community needs your input! [Share your idea] 24 other people contributed today.

Autonomous orientation (A)	**Impersonal orientation** (I)	**Controlled orientation** (C)
Make our community a better place! [Share your idea] 24 other people contributed today. Your ideas matter.	Change may be beyond our control... [Share your idea] but there's a chance someone sees your idea and considers it.	Become a Contributor of the Month! [Share your idea] $20 gift cards for Top 5 contributors. Announcement on official website.

To test participants' volunteering motivation, I used the Motivation to Volunteer Scale (MVS) questionnaire developed by Grano et al. [14]. It contains 24 prompts with ratings on a scale of 1 to 5 relating to the six types of motivation/regulation identified by the SDT. I also asked participants about how often they participated in volunteering activities in the past three months on average, on a three-point scale (never, less than once a week, about once a week or more).
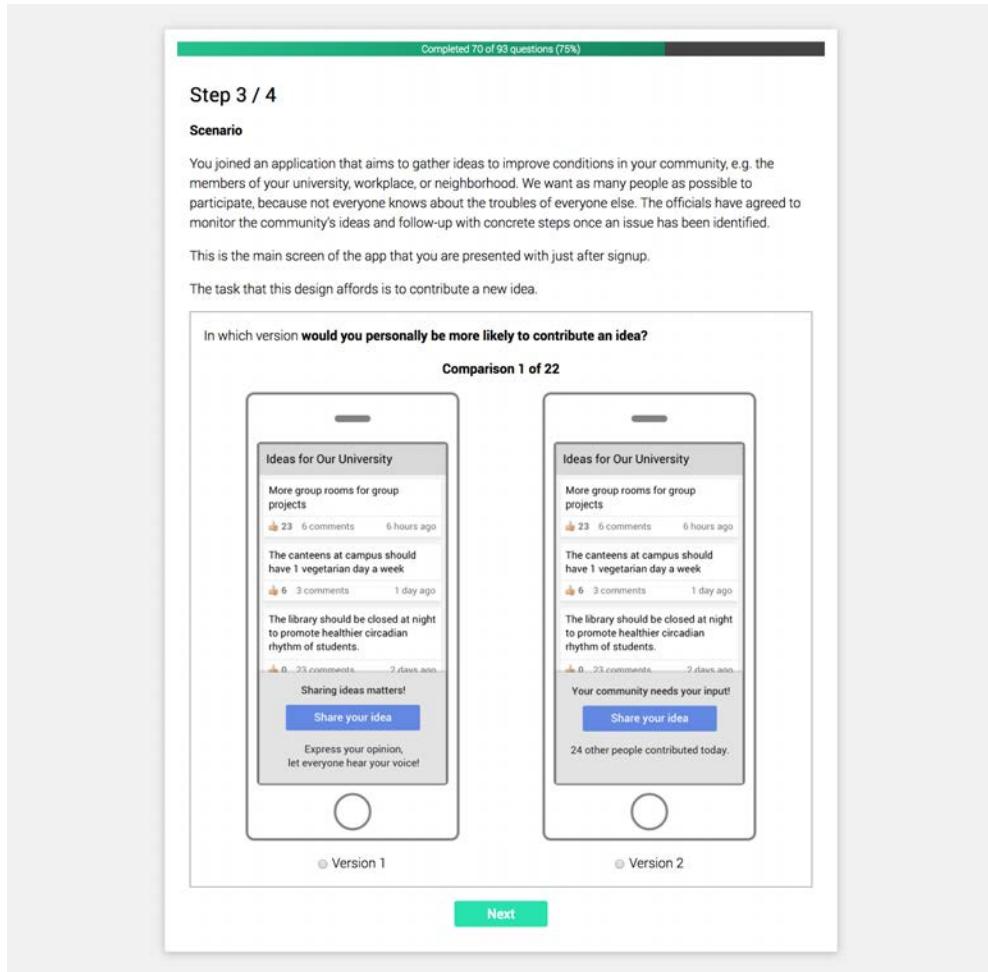
As is standard practice in conducting research experiments on MTurk, I used a number of measures for quality control [38]. I used a set of trapping questions to check participants' attention. Both the GCOS and MVS questionnaires included bogus items requiring participants to answer with a certain rating. The pairwise comparison task included one repeated measure to check for consistency. I also measured the time-to-complete to detect outliers.

The survey was implemented using custom code and the Flask[1] web framework. Two screens can be seen in Figure 4.2.

---

[1]   http://flask.pocoo.org/

(a) Survey start page.

(b) Pairwise comparison step.

Figure 4.2: Two screens from the custom-built survey website. The other steps are depicted in Appendix B.1.

4.4 Data Collection

I collected responses from 150 MTurk users. For completing the task, which was estimated to take up to 20 minutes, they were rewarded USD 3.50 (for a nominal hourly wage of USD 10.50). As the survey involves a considerable amount of reading, the experiment was restricted to MTurk users from the U.S. to increase the chance of participants being able to communicate in English fluently. I also set a minimum Approval Rate of 90%. After eliminating responses that failed the attention checks or that were impossibly quick (time-to-complete of less than 5 minutes), 99 responses remained.

Of the participants, 32% were female (slightly less than a usual U.S.-based MTurk distribution [31]). Attained education and age were uniformly distributed over the complete range. The distribution of demographic groups is depicted in Figure 4.3.
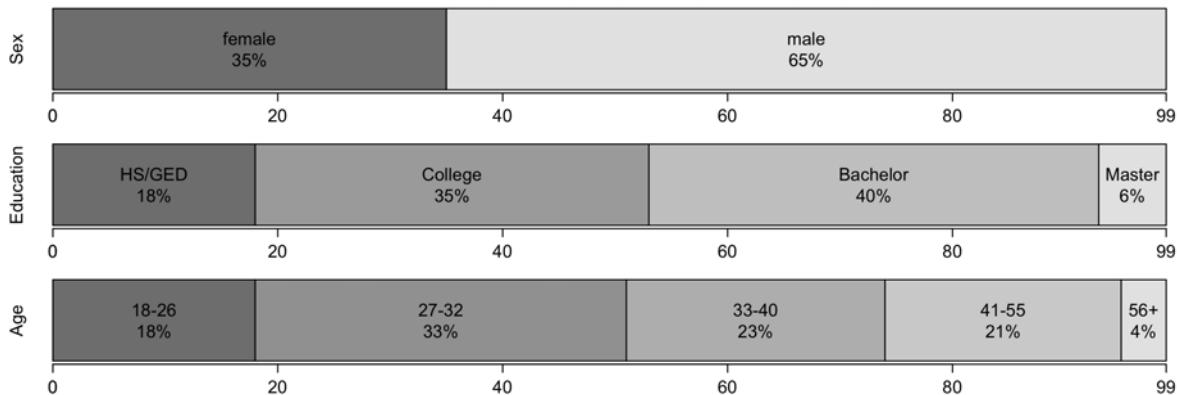

Figure 4.3: Demographic data for Study 1 (N=99). The data is mostly uniformly distributed as expected from the MTurk population [31].

4.5 Results

4.5.1 Test scores

The distribution of the all subscale scores (Figure 4.4) follows the trend of data reported in previous work [8, 14], indicating validity of the results obtained from MTurk workers. In terms of GCOS, most participants scored highest on the Autonomy scale, followed by Control and Impersonal. The three scales are nearly normally distributed, with the Autonomy scores being skewed towards the upper end of the scale. As for MVS, participants tended towards the autonomous side of the spectrum (high scores on the identified, integrated, intrinsic subscales), but exhibited a broad range in all of the subscales.
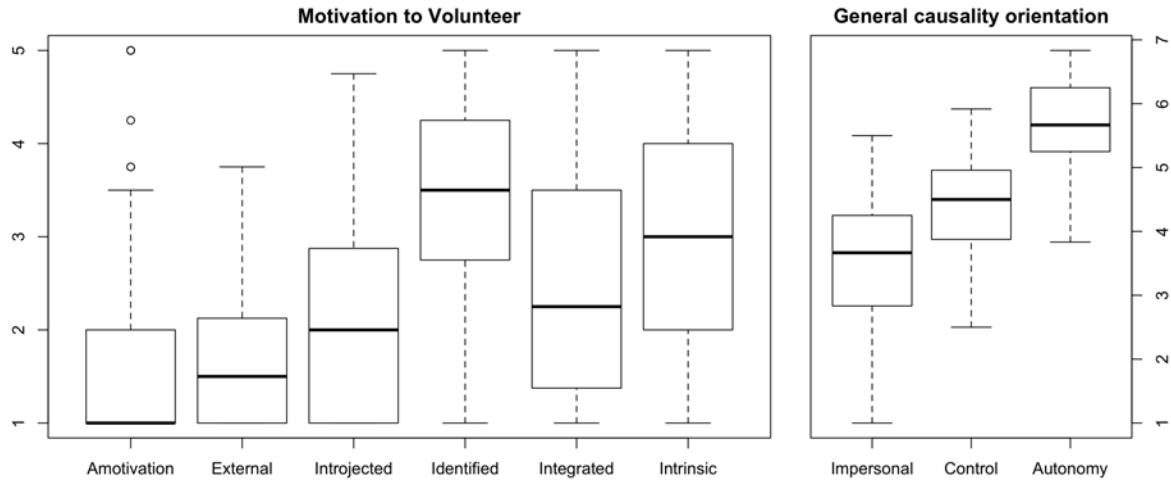
Figure 4.4: Distribution of MVS and GCOS scores (N=99). The distribution of the scores follows the trend of data reported in previous work [8, 14], indicating validity of the results obtained from MTurk workers.

4.5.2 H1: Diverse Overall Preferences

Looking at only the individually top-ranked preferred version, my hypothesis H1 is supported, as there was no consensus on the "best" version. In total, 65 (63.1%) participants preferred version C and 14 (13.6%) preferred version A. The baseline (BL) and Impersonal-oriented (I) versions consistently ranked lowest and the other versions were somewhere in between. Despite other versions also being favored by some people, based on these results, one might conclude that offering extrinsic rewards like in version C is generally preferable.

4.5.3 H2: Individual Differences

To answer if there were individual differences in preferences (H2), I constructed a Loglinear Bradley-Terry model (following the examples of [16]) using the results of the pairwise preferences and subscales scores as subject covariates. The model predicts ratings for the different versions based on specified factors. I created new factors for each test scale by binning scores into "low" and "high" groups corresponding to the lower and upper quartiles of the recorded scores.

The model estimates support my hypothesis H2, with the most significant differences between the A and C versions (Figure 4.5). Regarding GCOS, participants with a low Impersonal score preferred $A > C > nC > nR > nA$, whereas participants with a high score preferred $C > A > nA > nR > nC$. As expected from the theory, the Autonomy scale showed significant effects towards the opposite direction (low: $C > A > nC > nR > nA$, high: $A > C > nC > nA > nR$). The Control orientation scale also had a strong correlation with preference for version C. Regarding MVS, the two scales at the end of the spectrum had the strongest effects (Amotivation score low: $A > C > nC > nR > nA$, and high: $C > A > nA > nC > nR$; Intrinsic score low: $C > A > nA > nC > nR$, and high: $C > A > nR > nC > nA$). These results lead us to the following claims:
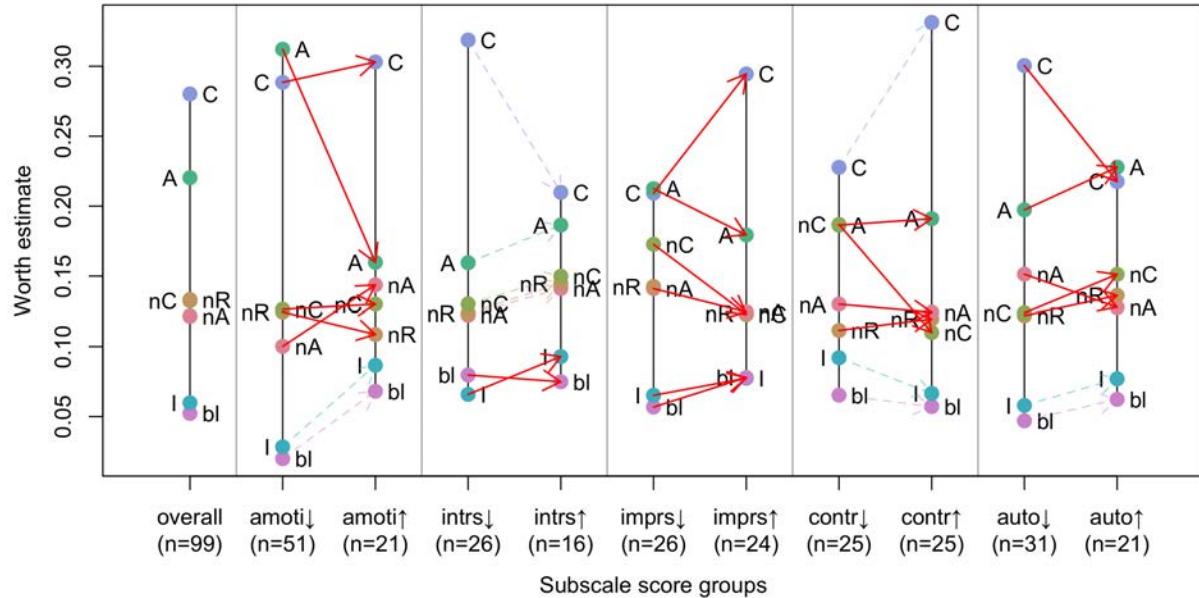
Figure 4.5: Version preference estimates of Bradley-Terry model of pairwise comparison data, overall and for groups of low↓ or high↑ scores on the subscales *Amotivation, Intrinsic, Impersonal, Control, Autonomy*. Solid red arrows indicate rank changes. Omitted subscales showed no significant differences.

- External control in the form of rewards (C) is generally preferred, however for users with high Autonomy-orientation, low Control-orientation, low Impersonal-orientation, or low Amotivation scores, an autonomy-supportive (A) design may be preferable.

- If external rewards (C) are not possible, supporting all three needs for autonomous motivation (A) is generally preferable over supporting only a subset of needs.

- If supporting all three needs is impossible, Competence-need (nC) or Relatedness-need (nR) supportive messages may be more effective for users with low Impersonal, low Amotivation, or high Autonomy scores, whereas Autonomy-need (nA) supportive messages may be more effective for the opposite cases.

4.5.4 Qualitative Data

The free-form explanations given by the participants serve as further evidence for my hypotheses: people cited various reasons for liking different versions, and the reasons were aligned with my expectation. Participants who preferred version A mentioned reasons such as "*It looks more friendly*" and "*I like where it reminds people that their contribution will help the community.*" Participants who preferred version nR said, "*I appreciated the sort of bandwagon/group mentality. Knowing how many other people have shared ideas would make me less reticent about expressing an idea.*"

Even though only a few people preferred the baseline and version I, they made sense of their choice: "*It's very simple and it doesn't insult the person contributing ideas by talking down to them*" (version BL) and "*I prefer a statement that doesn't try to make me feel guilty for not sharing an idea*" (version I). These

versions might only appeal to a minority of users, but it could be worth investigating supporting them, as well.

4.6 Limitations

Results from Study 1 have a number of limitations. First and foremost, the preferences are all self-reported. Asking people what they think they will do in a hypothetical scenario and observing actual behavior can yield very different results, an effect also known as *hypothetical bias* [36]. The MTurk population is also possibly non-representative for the general population, as they are paid workers. Additionally, this study only looked at a single component of a hypothetical application instead of a complete integrated environment.

To address these issues, I designed a real-world application based on these results and conducted a field study, explained in the next chapters of this thesis.

5

Design of a Research Probe

Having identified possible interventions in my first study, I proceeded to design an application to use as an apparatus in my field study, study 2. The mobile web application used for this study was designed to allow for voluntary participation of the crowd regarding social matters of a community. Some of its main screens are depicted in Figure 5.1. I argue that this app can be seen as a basic example of a crowd-civic system.
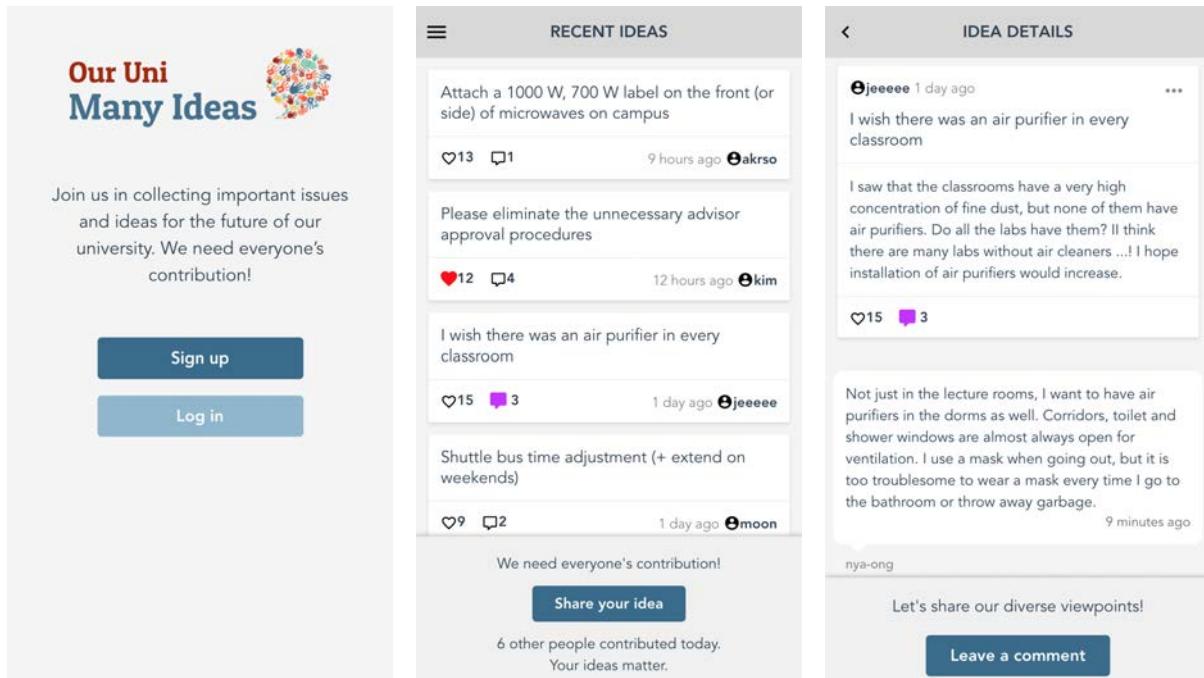
5.1 Design Goals

Using results from the first study and drawing from previous work on crowd reporting systems [53] and online community design [29], I identified the following minimum requirements:

- Shared, attainable, actionable goals with real impact

- Small but relevant individual contributions towards these shared goals

- Interaction between members of the crowd, for democratic deliberation and empowerment through large numbers

Additionally, I formulated the following design goals to make it feasible to run a short-term deployment study using this system:

- General and simple tasks that do not require special knowledge and enable meaningful interaction within a short period of time

- Open tasks that do not limit quantity and quality, to be able to measure differences in engagement

- General basic design that allows for implementation of different motivation-supportive design strategies on top

(a) Home page before login with branding and main goal of the app.

(b) Main feed of ideas with personalized call-to-action to participate.

(c) Idea detail view with ways to engage by commenting or showing support.

Figure 5.1: Case study application's main screens. Displayed contents are real examples. Real university name is replaced with a generic name.

5.2 Implementation

To satisfy these requirements and goals, I developed *Many Ideas*, a system supporting communities to report problems and generate ideas to improve conditions for its members. Users share their views through creating issues, discussing, or voting. The system is based mainly on an altruistic motivation to help "the greater good", to improve the lives of every member of that community.

Communities supported by this system can be workplaces, neighborhoods, or other kinds of shared spaces. For this study, I decided to deploy the system for the members of one university. As they often live, study, and work here, students, faculty, and staff spend a considerable amount of time in a shared environment, making it a prime target for this sort of application.

Many Ideas offers four key interactions: contributing your own idea, browsing ideas posted by others, commenting on ideas, and expressing support by pressing a heart button. Browsing is supported by a feed of ideas on the main screen after login.

The application was implemented as a responsive single-page static frontend using Vue.js and a backend server using the Django framework. To support open research and enable future experiments using this app as a basis, the source code is available on GitHub at http://github.com/graup/manyideas.

---

2   https://firebase.google.com/docs/hosting/
3   https://www.pythonanywhere.com/
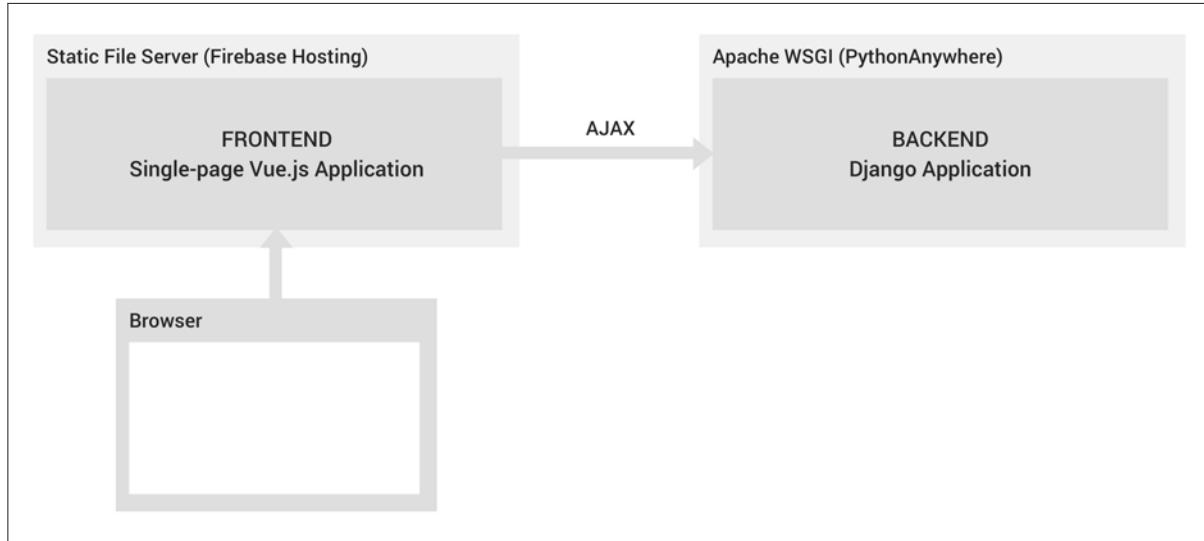
Figure 5.2: Application architecture overview. The user's browser retrieves the frontend from a static file server (Firebase Hosting[2]). The application talks to the backend (hosted on PythonAnywhere[3]) via AJAX.

5.2.1 Backend

The backend was implemented using the Django framework[4]. It consists of the data model and an API. There are two major sub-applications, *issues* and *experiment*. The *issues* application holds data models (Figure 5.4a) and API for user generated content such as issues (ideas) and comments. The *experiment* application holds data models (Figure 5.4b) and logic necessary for the treatment assignments. The assignment algorithm ensured well-balanced assignments according to the description in subsection 6.3.3.

For displaying the main feed of ideas, a custom ordering was implemented using an algorithm like Reddit's "What's hot" ranking [45]. It boosts popular items to the top for some time. This is to strike a balance between showing popular ideas first and upvote inflation, which would bury unpopular or new posts.

5.2.2 Frontend

The frontend is a responsive web application, designed to be used on mobile screens. It uses Vue.js[5] as the underlying framework. A diagram of the main views can be seen in Figure 5.3.

Signup and Onboarding. Before signing up, users are asked to read and accept the informed consent text as required by IRB. After providing username, email, and password, users proceed to the onboarding screen, which elicits additional information used for personalization as described in subsection 6.3.3.

Feed (Figure 5.1b). The main feed of recent and popular ideas. Clicking on an idea leads to the Idea Detail view. The feed view also contains a prominent call-to-action component.

[4]   https://www.djangoproject.com/
[5]   https://vuejs.org/
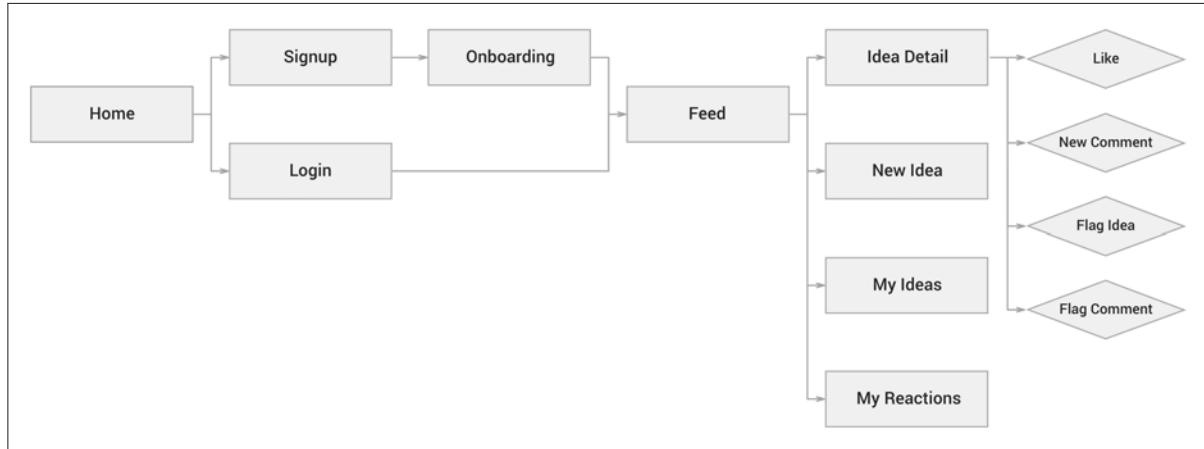
Figure 5.3: Frontend view flow chart

Idea Detail (Figure 5.1c). Users can see the description text of the idea as well as any comments. They can also press the Heart button to "like" this idea as well as add their own comment.

Flagging/reporting. Users can flag both ideas and comments for inappropriate content. When content is flagged, a notification is sent to an administrator who can review the report and decide to delete the infringing item.
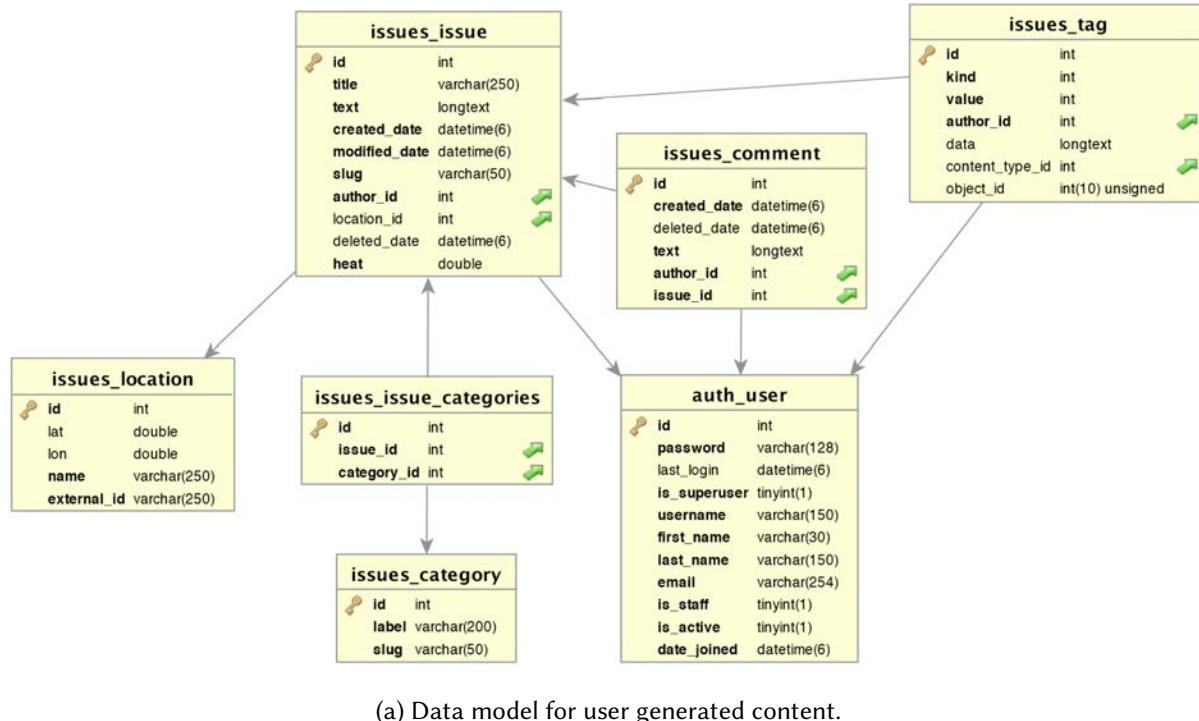
New Idea. When adding a new idea, users are prompted for a title and a description, and optionally can choose a physical location this idea relates to. This feature uses the HTML5 Geolocation API as well as the Overpass Turbo API[6] which is used to query places of interest from OpenStreetMap[7] near the user's current location.

My Reactions. This view shows all ideas that the user has interacted with, either by commenting or pressing the Like button. This feature was added in response to feedback during beta testing.
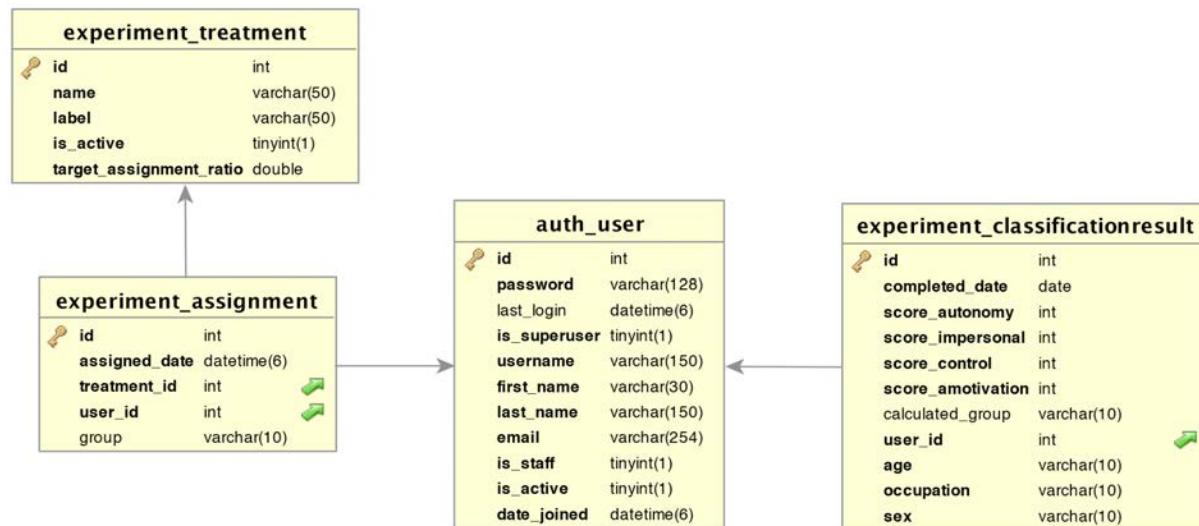
Apart from recording user-generated content, additional interaction data was collected using Google Analytics. On each navigation event and click on a button, an event was saved along with the user id.

---

[6]   http://overpass-turbo.eu/
[7]   https://www.openstreetmap.org/

(a) Data model for user generated content.

(b) Data model for experiment treatments and assignments.

Figure 5.4: Entity relationship diagrams. Generated with DBVisualizer.

6

Study 2: Field Study

To address the shortcomings of the first study, I ran a real-world campaign to measure effects of designs based on the Self-Determination Theory (SDT) on volunteering behavior. In an open call, members (students, faculty, and staff) of a large technical university in Korea were invited to participate in a community of voluntary participation to gather ideas to improve conditions on campus, following the design introduced in chapter 5.

6.1 Hypotheses

To understand how engagement in civic participation applications can be affected by design based on SDT, I formulated the following hypotheses. In light of my previous results, I narrowed down the space of motivations to controlled and autonomous motivation.

H1: Both autonomous and controlled motivation strategies are effective in increasing one-time, short-term voluntary participation.

H2: Personalized motivational support is more effective for increasing participation than either strategy applied unconditionally.

H3: External regulation in the form of monetary rewards is effective in increasing one-time, short-term voluntary participation, but detrimental to extended civic engagement.

6.2 Method and Measures

6.2.1 Main procedure

I opted to do a real-world deployment study because of the voluntary nature of the task and my goal of measuring how different designs affect individual motivations. Any lab setting or direct financial

reward could undermine users' individual intrinsic motivations [44]. It is therefore important to create an ecologically valid setting.

Upon sign up, users were classified into one of two groups predicted to prefer one design over the over. They were then uniformly distributed between the three experiment conditions implementing different strategies based on SDT, as explained previously. In this setting, I collected several measures of engagement under different motivation-supportive conditions. The assignments were fixed per user (between-subjects design). This resulted in a 2x3 factor design.

The study spanned a period of 8 days. Some positive effects of motivation-supportive design might not manifest themselves immediately, so rather than collecting a snapshot of user data, I wanted to see how engagement varies over time. Fluctuating engagement is one of the core observations of previous experiments in this field [2].

To evaluate my hypotheses, I collected a number of interaction data during the deployment. I collected the number of posts, likes, comments, and events (including navigating pages, creating content, and clicking the heart button) made by each user in total and per session, the number of sessions, and time spent per session. I also analyzed the contents users contributed in terms of length. Upon signup, I collected information about sex, age, and status in the university.

6.2.2 Post-survey

To find out more about the individual differences between users and feelings towards continued voluntary participation, I distributed a questionnaire to every user after the end of the study period. Whereas no general compensation had been offered to users of the application, survey respondents were paid KRW 7000 (ca. EUR 5.50) for their participation.

The post-study survey included the questionnaires for general-causality orientation (GCOS [8]) and motivation to volunteer (MVS [14]) as in Study 1. I collected this data to compare the groups assigned during the deployment, which were based on limited information, with groups based on more complete data. Additionally, I asked for specific reasons why users decided to join the application or why they stopped using it. I also asked about the general usability of the app using the AttrakDiff mini questionnaire [15], and some open-form feedback. The exact questions can be found in Appendix A.2.

6.3 Data Collection

6.3.1 Recruitment Campaign

To recruit users, I ran a promotion campaign through multiple on-campus channels, including physical posters (Figure 6.1), and posts on mailing lists, school online communities, and social networks. The public campaign deliberately did not specify any rewards, to minimize the potential confounding of

(a) "Cafeteria menu, is this the best?"

(b) "General electives. Applied for 4, got 0. For real?"

(c) "The internet speed in Korea's best technical university..."
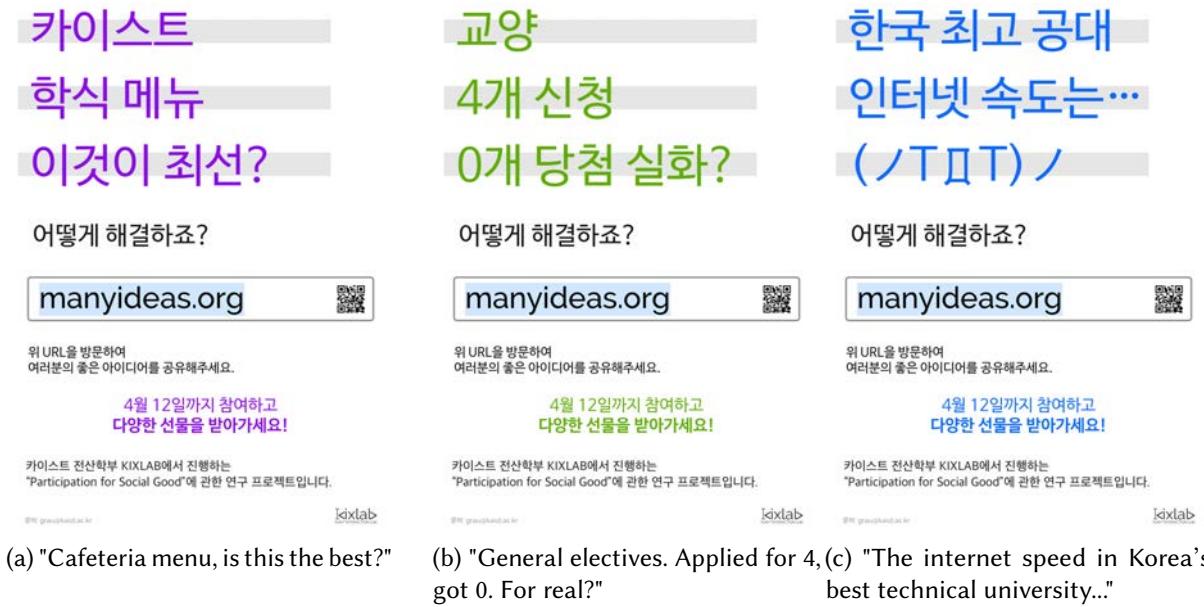
Figure 6.1: Poster used for deployment campaign with varying slogans. The text below explains the incentive event. The information is deliberately leaving out precise details.

intrinsic motivation. I tried to attract attention by using a number of catchy slogans and raise curiosity by including vague messages like "There are various rewards for participation. Sign up to find out more!" After the first 3 days, I also sent an email reminder to every user so far to encourage them to revisit the app. On all promotional material, I indicated that the incentive period will conclude at a specific date (8 days after launching the study).

6.3.2 Experiment Conditions

In light of my previous results, I narrowed down the space of supported motivations to controlled and autonomous motivation. In Study 1, these two conditions showed the biggest differences. Thus, I designed three different message strategies for the application: a baseline, an autonomous-motivation supportive strategy, and a controlled-motivation supportive strategy.

I envisioned two distinct tangible rewards for users, reflecting the internal and external regulatory styles associated with these kinds of motivations:

- *Idea curation*: For the autonomy-supportive treatment, I offered to curate three "promising ideas" that I will publicize to the whole school and also take active steps towards their implementation.

- *Cash prize*: For the control-supportive treatment, I offered a raffle of 10x KRW 20,000 (ca. EUR 15.80) cash prizes among all active users.

The rewards were chosen in accordance with SDT. The *cash prize* is an external reward tied to a single user. It does not create an explicit connection with the amount and quality of a user's participation. Everyone was eligible, as long as they interacted in any way. On the other hand, the *idea curation*

All versions		
What is your idea to make (university name) a better place? It could be about facilities, organization, social problems, …		
Baseline	**Autonomy-supportive** (Idea curation)	**Control-supportive** (Cash prize)
(No extra text)	**Let's have a real impact together!** Among all ideas submitted until April 12th, we will hand-pick three promising ideas, present them to the whole school, and follow up with concrete steps to support their implementation!	**Participation reward:** Among all contributors until April 12th, we will randomly select 10 members to win $20.

Table 6.1: Example for version differences in Study 2. These messages are shown on the top of the *New Idea* view. The baseline only contains the absolutely necessary information, while the Autonomy- and Control-supportive versions include references to the respective incentives and regulatory styles. The complete collection of messages can be found in Appendix B.3.

reward is related to intrinsic motivations and identified or integrated goals, such as real-world impact and recognition of the importance of one's contribution. Any tangible reward in that condition is not something that the user receives directly, but rather something their idea, or effectively the community, benefits from.

Both rewards were promoted in-app according to the treatment a user was assigned to, i.e. users were only aware of one of the rewards. For the baseline treatment, a section on the app's FAQ page explained that there were rewards for which everyone was eligible, but that I wouldn't disclose them yet. The two other treatment conditions also contained an explanation that even though there might be different rewards, every user will unconditionally be eligible for all rewards.

As another treatment, I considered an Amotivation-supportive strategy similar to the one from Study 1 (empathizing with user's feelings of incompetence or belief that achievements are beyond their control). Since the number of users was not likely to gain meaningful results for this minority treatment, I decided to leave this option for future, larger-scale deployments.

The two strategies, supporting either controlled or autonomous motivation, were added to the design of my research probe. I selected 6 key elements from the application that I created alternative messages for, relating to support of the different motivations. One example for the version differences can be seen in Table 6.1. I aimed to keep these treatments simple and salient to maintain the connection to the theory and to not confound different motivational processes as much as possible.

All users were Korean native speakers, and the complete application including all prompts and motivational messages was translated to Korean by a group of native speakers under my guidance. Users contributed content in Korean.

6.3.3 User Classification for Personalization

For personalization, some personal data is necessary. As I did not want to change treatments during the experiment to avoid confusion, we decided to classify users into groups that are likely to be affected by autonomous or controlled motivation based on information gathered at the time of registration. However, given the voluntary and quick-use nature of the app, I did not want to subject users to completing the complete personality questionnaires that we used in Study 1, due to a fear of causing annoyance and early dropouts.

To reduce the number of questions necessary to prompt users, I inferred a decision tree from the data collected in Study 1. The raw responses to the items of questionnaires served as independent variables and the relative preference of the autonomy-supportive version over the control-supportive version as the dependent variable. As the distribution between the two preference was very imbalanced (1:3), we used the method of Randomly Oversampling Examples [34] before partitioning the data using a Conditional Inference Tree (Figure 6.2). Using k-fold cross validation resulted in an accuracy of 70% on a generated tree containing only four questions of the original 63 questions. These four questions (Figure 6.3) were one each from the personality subscales Amotivation (from MVS), Control, Impersonal, and Autonomy (all from GCOS). The decision nodes were all intuitive (e.g., people with a high Amotivation score were classified as preferring the control-supportive version), so I decided to use this classifier. I will evaluate and discuss this choice in depth later.
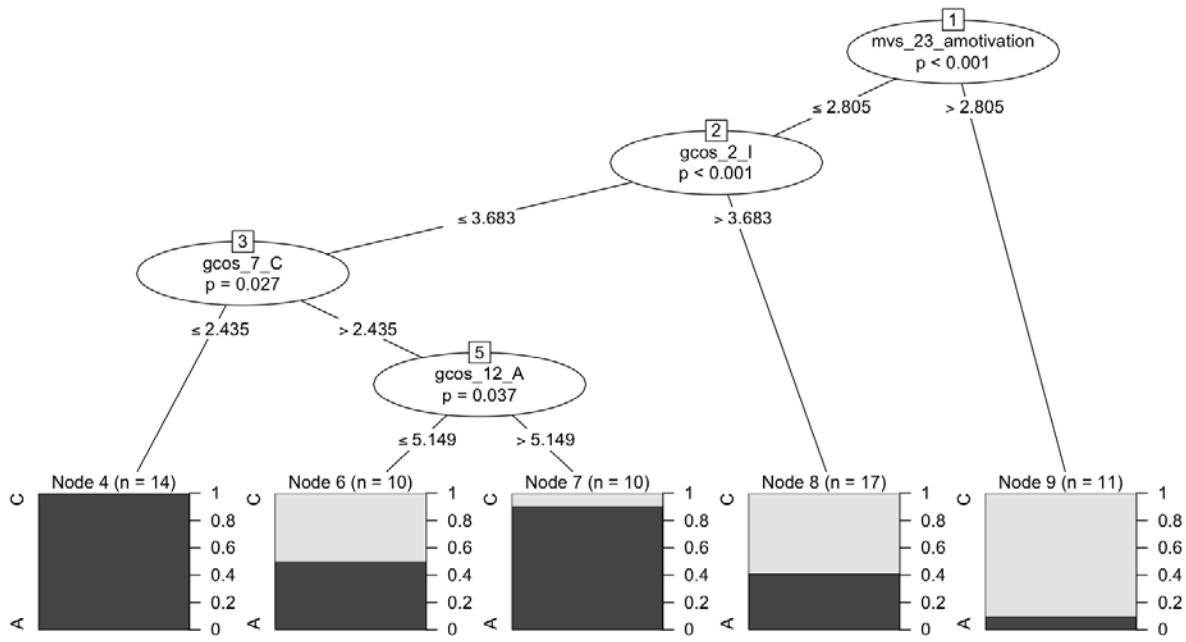
Figure 6.2: Conditional inference tree for preference of A over C version based on data from Study 1. The exact text for the selected prompts is shown in Figure 6.3.

To help us better understand your personal attitude towards civic participation, please answer the following questions.

Imagine the following situations and how you would respond.

(1) You have a school-age daughter. On parents' night the teacher tells you that your daughter is doing poorly and doesn't seem involved in the work.

    You are likely to scold her and hope she does better.

(2) You have been invited to a large party where you know very few people. As you look forward to the evening, you would likely expect that:

    You'll try to fit in with whatever is happening in order to have a good time and not look bad.

(3) A woman who works for you has generally done an adequate job. However, for the past two weeks her work has not been up to par and she appears to be less actively interested in her work. Your reaction is likely to be:

    Ask her about the problem and let her know you are available to help work it out.

(4) Whether you volunteer or not, what do you think is the reason for that?

    I don't know; I can't see what I'm getting out of it.

Figure 6.3: Classification instructions and questionnaire presented to users directly after signup. All questions used a 1-7 rating. The questions relate to the following scales: 1 - Impersonal, 2 - Control, 3 - Autonomy, 4 - Amotivation. Questions 1-3 are from GCOS [8], question 4 from MVS [14].

6.4 Results

6.4.1 Participants

Within 8 days of the live deployment, 120 users signed up for my application. Among the participants, 94% were undergraduate or graduate students, the rest faculty and staff. There were less women than men, 29%, however, this is higher than the overall 20% ratio of women in this university. The distribution of demographic data is depicted in Figure 6.4.

Users posted 72 ideas (0.6 per users), 62 comments (0.5 per user), and expressed their support using the heart button 357 times (3.0 per user). The majority, 79 users (66%), performed at least one of these interactions. In total, I recorded 528 unique visitors (visit-to-signup ratio of 23%). To track the source of traffic, I provided different URLs for different channels. Approximately half of the traffic came from my posts to the campus-internal discussion forum, 25% through Facebook shares, 10% through the posters, and the rest accessed the app directly, e.g. through word-of-mouth.

I observed continuous engagement in the application, with spikes whenever I made another promotion post on any of the channels. Figure 6.5 shows the continuous trend of number of visitors, signups, and new ideas during the study period.

Upon signup, users were assigned to groups and treatment conditions using the result from four personality questions and a pre-computed inference tree, as described before. My model assigned 59% of the users to the "controlled" group and 41% to the "autonomous" group. Within each group, exactly 1/3 of the users were assigned to each of the three treatments.

I collected post-survey responses from 38 users (who had been assigned to these study treatments: 14 autonomy, 11 baseline, 13 control), a response rate of 32%.
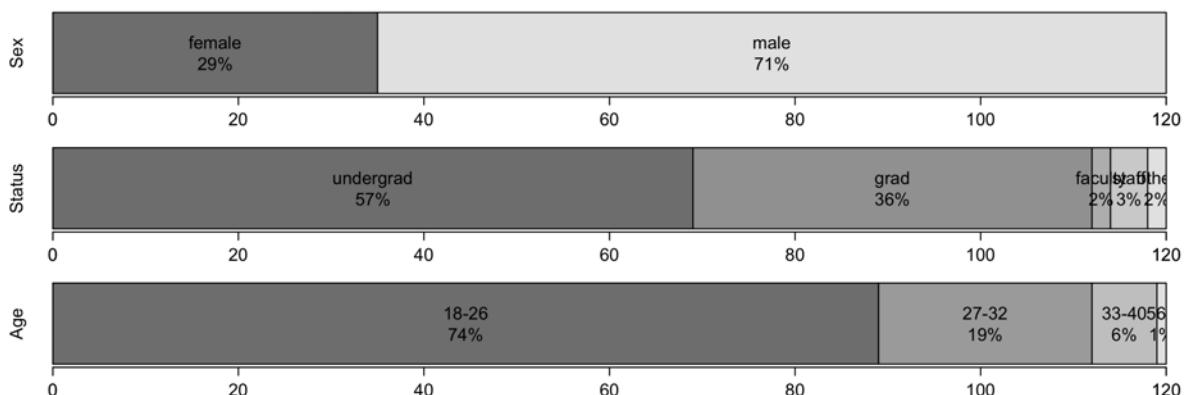

Figure 6.4: Demographic data for Study 2 (N=120)
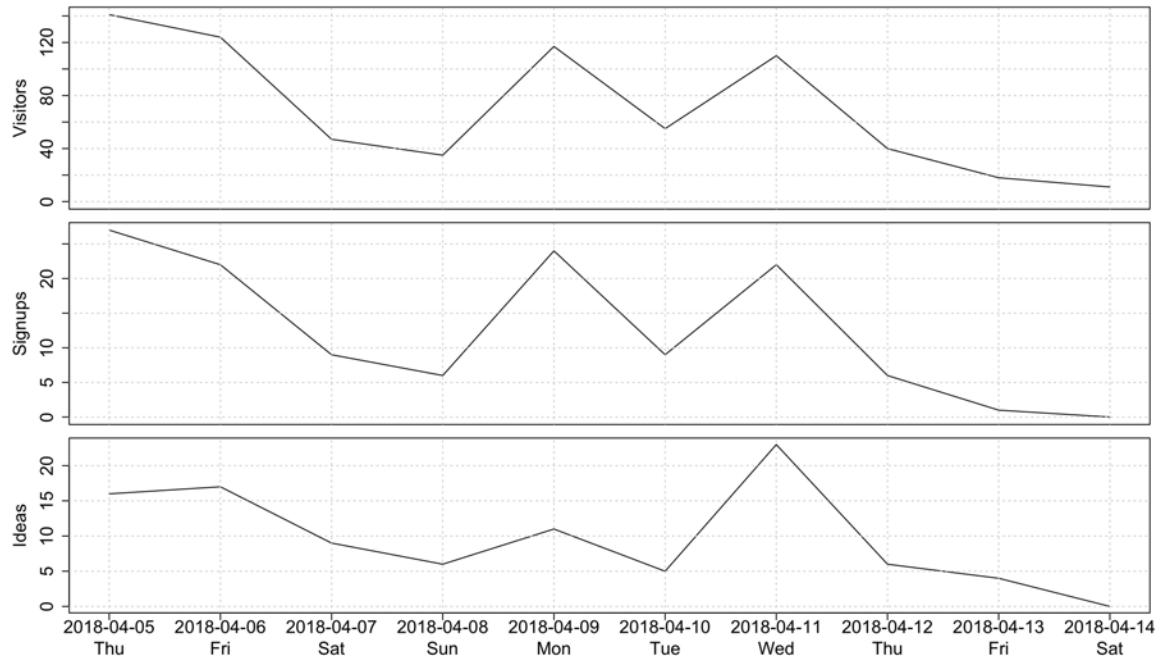
Figure 6.5: Number of visitors, signups, and new ideas per day during the study period. The peaks coincided with repeated posts to the university's online community.

6.4.2 Post-hoc Clustering

I used the post-hoc personality data from the survey to cluster respondents into two groups using k-means (Figure 6.6, Table 6.2). The clusters found align with theory intuition: one cluster has participants with relatively high scores on scales related to autonomous motivation, the other on scales related to control or amotivation. I found scores from the MVS scales to be more descriptive than GCOS scores: applying k-means to only MVS data yielded the same solution as applying it to both constructs combined, whereas using only GCOS data yielded no meaningful solutions. Comparing the original groups to the post-hoc ones yielded an accuracy of only 44%, indicating a bad fit of my original classification model.

6.4.3 H1: General Effectiveness of Control and Autonomous Motivation Strategies

The monetary reward that was offered in the controlled-motivation condition was clearly successful. Users in that condition showed higher engagement compared to the two other conditions. They posted more ideas (posts per user: Control 0.94, Baseline 0.41, Autonomy 0.46; ANOVA $p<0.01$; post-hoc multiple comparisons with Tukey method $p<0.05$ for Control treatment compared to both other treatments; GLM for Poisson distributed count data) and wrote longer posts (average characters per post: Autonomy 36, Baseline 33, Control 64; n.s.). Differences between the autonomy condition and baseline were also positive, but much weaker and non-significant. This means that hypothesis H1 is partially supported.
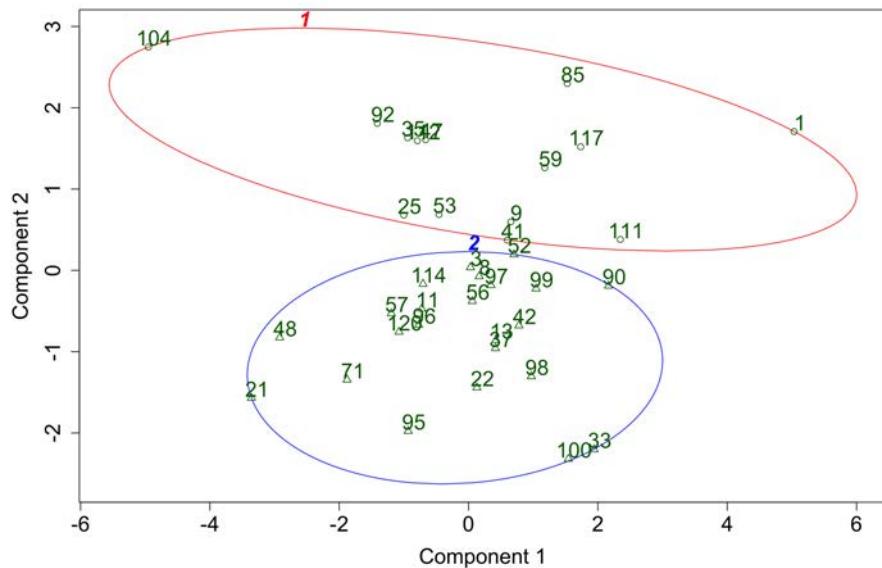
Figure 6.6: Kmeans cluster solution of MVS scores from post-survey. Plot generated with the `clusplot` method from R's `cluster` package. These two principal components explain 76.59% of the point variability.

Scale	Cl. 1	Cl. 2
Intrinsic	3.27	*3.54*
Integrated	3.10	*3.15*
Identified	3.04	*3.76*
Introjected	*3.35*	2.33
External	*3.27*	2.40
Amotivation	*3.55*	1.99

Table 6.2: Cluster means of Kmeans cluster solution of MVS scores from post-survey. The respective higher value is *highlighted*. The solution seems reasonable: higher scores in subscales related to more controlled motivations are in cluster 1, while higher scores in subscales related to more autonomous motivations are in cluster 2.
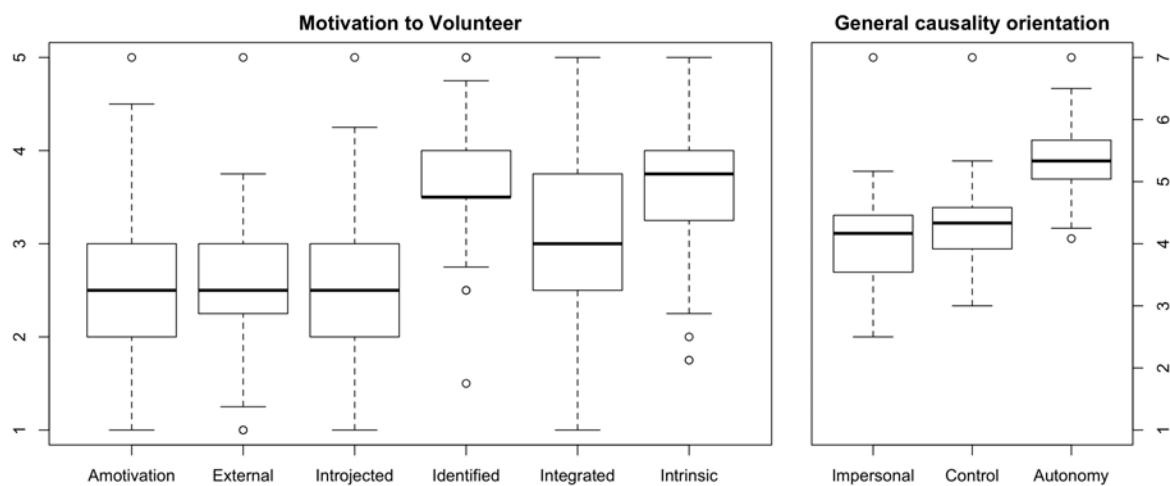
Figure 6.7: Distribution of MVS and GCOS scores of survey respondents (N=37). The data follows a similar trend as the data from Study 1 (see Figure 4.4).

6.4.4 H2: Interaction Effects of Treatment and Group

There were no significant interaction effects between treatments and user groups as classified from the signup questionnaire. None of the collected engagement measures showed statistically significant differences. It is likely that the classification based on four questions was inaccurate.

However, Hypothesis 2 is supported when taking post-hoc data into account. For survey respondents, I used the post-hoc groups instead of the ones assigned during the deployment for further analysis.

Being in the treatment condition that supported the individual motivation orientation had positive effects on the number of interactions (Figure 6.8), whereas being in the opposite group showed negative effects (ANOVA for number of interactions p<0.01 for treatment, group, and interaction; GLM for Poisson distributed count data). Compared to each group's baseline mean, for control-motivated participants, I counted 32% more interaction events when in the control treatment and 30% less events when in the autonomy treatment (Autonomy: 30, Baseline: 43, Control: 57; multiple comparisons with Tukey method, all p<0.01). For autonomy-motivated participants, I counted 2% more events when in the autonomy treatment and 8% less events when in the control treatment (Autonomy: 40, Baseline: 41, Control: 38; multiple comparisons n.s.).

Though not significant, other collected measures show similar trends. For example, users in the respective motivation-supportive treatment contributed ideas with more content (median baseline for both groups: 54 characters per post; control group in control treatment +93, in autonomy treatment -54; autonomy group in autonomy treatment +27, in control treatment +6). The overall data suggests that SDT-related personality scores are correlated with actual behavior, but should be regarded with caution due the low number of responses.
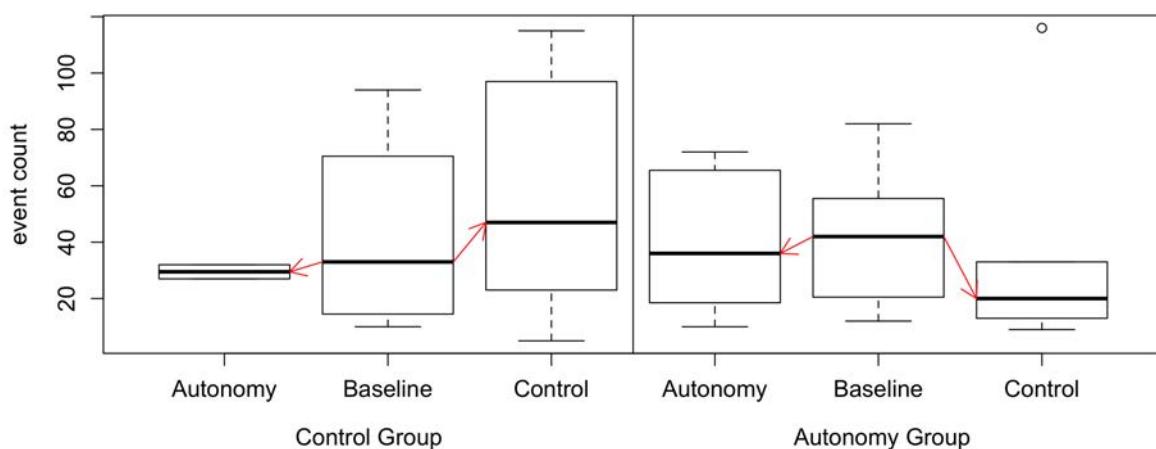

Figure 6.8: Number of interactions by treatment and post-hoc group. Motivation-supportive treatments resulted in an increase of interactions and vice-versa. Omnibus test (ANOVA): p<0.01. Multiple comparisons (Tukey method): for control-motivated group (left half) all p<0.01, for autonomous-motivated group (right half) n.s.

Table 6.3: Engagement data by treatments. Mean and standard deviation. Left table shows data for all users, right table only for those who responded to the survey.

	All users			Only survey respondents		
Treatment	Autonomy	Baseline	Control	Autonomy	Baseline	Control
N	38	35	32	10	11	10
event count	27.2 (23.7)	34.0 (27.2)	30.5 (32.8)	38.1 (22.5)	41.6 (28.7)	47.8 (44.4)
idea count	0.5 (1.1)	0.5 (1.0)	1.0 (1.9)	0.9 (1.9)	1.0 (1.3)	2.2 (2.9)
like count	2.2 (3.2)	3.6 (5.2)	3.3 (6.8)	2.5 (2.2)	5.9 (6.4)	7.9 (10.8)
comment count	0.4 (1.3)	0.6 (1.1)	0.8 (2.1)	1.2 (2.3)	1.0 (1.3)	1.3 (2.7)
use duration (min)	10.4 (19.4)	11.8 (14.5)	12.0 (21.3)	13.2 (15.9)	12.3 (13.7)	17.2 (34.6)
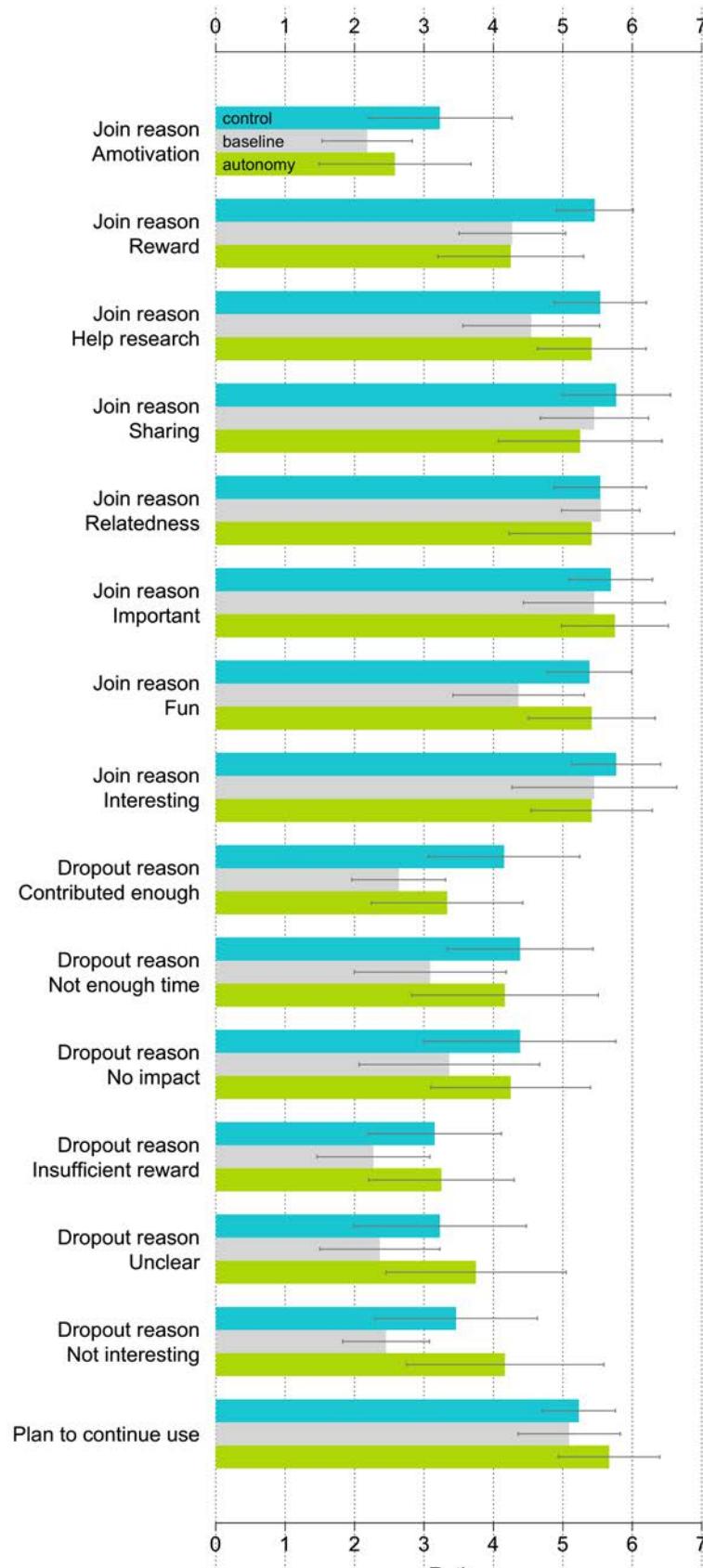
Table 6.4: Engagement data by post-hoc groups (only for survey respondents) and treatments. Mean and standard deviation.

Group	Control			Autonomy		
Treatment	Autonomy	Baseline	Control	Autonomy	Baseline	Control
N	4	4	6	8	7	7
event count	29.5 (3.5)	42.5 (37.8)	57.4 (47.2)	40.2 (24.9)	41.1 (25.6)	38.2 (44.4)
idea count	0.0 (0.0)	1.2 (1.9)	1.6 (1.3)	1.1 (2.0)	0.9 (0.9)	2.8 (4.1)
like count	2.0 (1.4)	4.8 (3.6)	9.2 (8.3)	2.6 (2.4)	6.6 (7.8)	6.6 (13.7)
comment count	1.5 (2.1)	1.0 (2.0)	2.6 (3.4)	1.1 (2.5)	1.0 (0.8)	0.0 (0.0)
use duration (min)	4.4 (6.3)	9.3 (14.3)	9.6 (11.6)	15.4 (17.0)	14.0 (14.1)	24.8 (49.2)
idea text length	0.0 (0.0)	38.2 (49.5)	133.0 (76.9)	73.2 (83.6)	120.7 (169.6)	146.2 (182.9)

Table 6.5: Survey responses by post-hoc groups (only for survey respondents) and treatments. Mean and standard deviation.

Group	Control			Autonomy		
Treatment	Autonomy	Baseline	Control	Autonomy	Baseline	Control
N	4	4	6	8	7	7
Plan to continue use	5.0 (1.6)	5.2 (1.0)	5.2 (0.8)	6.0 (0.8)	5.0 (1.4)	5.3 (1.0)
Dropout reason Not interesting	4.2 (2.2)	2.5 (0.6)	4.2 (1.9)	4.1 (2.4)	2.4 (1.3)	2.9 (1.7)
Dropout reason Unclear	5.5 (1.3)	3.2 (1.5)	3.7 (2.1)	2.9 (1.8)	1.9 (1.2)	2.9 (2.0)
Dropout reason Insufficient reward	3.8 (2.2)	2.0 (0.8)	3.8 (1.6)	3.0 (1.4)	2.4 (1.6)	2.6 (1.3)
Dropout reason No impact	4.8 (2.1)	4.2 (1.7)	4.0 (2.1)	4.0 (1.8)	2.9 (2.3)	4.7 (2.4)
Dropout reason Not enough time	6.0 (0.8)	3.2 (2.1)	3.8 (1.7)	3.2 (2.0)	3.0 (1.8)	4.9 (1.6)
Dropout reason Contributed enough	3.2 (2.5)	2.8 (1.0)	3.5 (2.2)	3.4 (1.4)	2.6 (1.3)	4.7 (1.1)
Join reason Interesting	5.8 (1.0)	4.2 (2.8)	5.3 (1.0)	5.2 (1.6)	6.1 (1.1)	6.1 (0.9)
Join reason Fun	5.0 (1.8)	3.8 (2.2)	5.0 (0.6)	5.6 (1.3)	4.7 (1.1)	5.7 (1.1)
Join reason Important	6.0 (0.8)	4.5 (2.4)	5.2 (1.0)	5.6 (1.4)	6.0 (1.0)	6.1 (0.7)
Join reason Relatedness	6.0 (1.2)	5.2 (1.0)	5.0 (1.3)	5.1 (2.2)	5.7 (1.0)	6.0 (0.6)
Join reason Sharing	4.2 (2.2)	5.2 (1.5)	5.7 (0.8)	5.8 (1.6)	5.6 (1.3)	5.9 (1.6)
Join reason Help research	5.2 (1.7)	4.2 (2.4)	5.7 (1.2)	5.5 (1.1)	4.7 (1.3)	5.4 (1.0)
Join reason Reward	4.8 (2.1)	3.8 (1.5)	5.5 (1.0)	4.0 (1.5)	4.6 (1.1)	5.4 (0.8)
Join reason motivation	3.2 (2.5)	2.8 (1.0)	4.2 (1.6)	2.2 (1.3)	1.9 (1.1)	2.4 (1.3)

Figure 6.9: Survey responses by treatment condition. Bars show mean values and double standard errors.

6.4.5 H3: Detrimental Effects of Controlled Regulation

Engagement data collect during the study does not support the hypothesis that controlled regulation has general detrimental effects on participation. All differences were either not statistically significant or towards the opposite direction. At least in the short term, the reward from the Control treatment seems to have nothing but increased participation.

Taking into account data collected during the post-survey paints a different picture. Figure 6.9 shows the survey responses to the join and dropout reasons. It can be seen that users in the Control treatment rated the reason "I think I have contributed enough" higher than other treatment groups. They also rated "I plan to continue to use this app" lower than others. Because of the low number of survey responses, differences are not statistically significant.

6.4.6 Reasons for Signup and Dropout

Participants reported various reasons why they chose to join the application, why they quit, and what they were hoping for. Reasons for joining and dropping out are depicted in Figure 6.9, grouped by the treatment the user had receive before answering the survey. There are no statistically significant differences, but it seems that the treatment might have had an effect. For example, users in the Control treatment rated the reward as being a important reason for joining the application — even though at the time of joining, all users had the same information. It seems that the treatments have changed participants' retrospective judgment. In general, both treatment groups rated all dropout reasons higher than the baseline group.

For completeness, the survey responses grouped by both treatment and post-hoc group are reported in Table 6.5. Due to the low number of respondents per group, I refrain from drawing any conclusions from this data.

Several people mentioned they were attracted by the opportunity to contribute to real change. Some said they found it interesting to see what others think. One person mentioned that the app could fill a void by offering a place to publicly express ideas and be credited for them. Another person highlighted the importance of listening to diverse opinions, even though they might seem "strange and stupid". A similar sentiment was visible in another respondent, who said it was "interesting to see others' ideas, regardless of how feasible they are." These responses reveal diverse motivations.

Multiple participants mentioned they wanted to see actual evidence of the system's impact, for example updates from the team or feedback from officials. As highlighted in numerous previous research [53], this would be a key factor for ongoing success.

6.4.7 Usability of the App

A poor overall design might have stronger effects on participation than any of the treatment conditions. To rule this out, I asked survey respondents to rate the usability and attractiveness of the app. The results imply that usability issues are not likely to have been major source of confounds.

In terms of quality as measured using the AttrakDiff questionnaire [15], the application scored medium-highly, with an average pragmatic quality of 5.3 (std 0.8), hedonic quality of 4.5 (std 0.7), beauty of 5.2 (std 1.2), and goodness of 5.7 (std 1.0) on 7-point scales. There is clearly room for improvement in terms of attractiveness, but generally people liked the application and did not seem to have any major problems.

During the deployment, no major issues occurred. The application was available and working as intended 100% of the study period. A few participants made suggestions for minor future improvements.

6.4.8 Follow-up on Submitted Ideas

Ideas submitted by users were generally well written and covered a wide variety of topics. As part of the *Idea curation* incentive, I had promised to pick three promising ideas and push them forward. For this, I created a scoring scheme and asked one member of the undergraduate student council to give their independent judgment. Among the most popular and most highly rated ideas, I finally selected three ideas.

All submitted ideas were scored on a scale of 0 — 3 (low — very high) for these constructs:

- Originality: how original, unique, or new is this idea?
- Feasibility: how feasible would it be to implement this idea?
- Depth: did the author give arguments or evidence? Is there a good amount of text or choice of words?
- Discussion: did this idea lead people to post meaningful comments? Or, does it have the potential to do so?

Most ideas received the highest possible score in feasibility, but the lowest possible score in originality. In Table 6.6, I report all ideas whose scores stuck out or that were very popular (i.e. they had more than 10 likes).

The final curated ideas were: 1. *Make a day of vegetarian meals once a week*, 2. *Eliminate unnecessary advisor approval procedures*, 3. *Fine dust mask vending machine*. These were chosen to cover a range of different types of ideas (one about inclusiveness, one about administrative processes, one about facilities).

6.5 Limitations

This study improved upon the first study in terms of ecological validity, but there are still a number of limitations.

The studied population was quite homogeneous (mostly Korean students). SDT claims to be universal, and the trends of the collected personality test scores are aligned with other populations, so there is no reason to believe SDT is not applicable. Still, it is hard to generalize the exact parameters of this study to larger or different populations.

The number of versions and messages in this study was limited on purpose and the chosen incentives were very salient. The nature of the incentives might have overshadowed more subtle effects of motivation support. I tried to create a motivation-supportive environment, but it can be argued that just changing a number of messages and incentives is not enough to make a complete environment.

Even though I saw some revisiting users during the 8-day study period, this is likely not enough time to measure any real long-term effects. A longitudinal study using a fully maintained commercial application would be necessary to do that. Especially regarding H3, detrimental effects of controlled regulation, much more data is needed to draw any conclusions.

Producing good real life outcomes was not the main focus of this work, even though I tried to do my best. The ideas submitted were useful, but all of a certain kind. There are a lot of things the design could be improved on to make users produce better content, regardless of motivation support or personalization.

Title	Rater's judgment	L	C
Attach a 1000 W, 700 W label on the front (or side) of the microwave on campus		16	0
Fine dust mask vending machine	*original*, *in-depth*,	14	3
We need faster internet	*in-depth*,	14	2
Please switch to the garbage cans that you step on or press the garbage cans (especially food items) in stores.		14	0
Shuttle bus to Endless Road		14	0
Eliminate unnecessary advisor approval procedures	*creates discussion*,	13	1
Let's make an electronic student ID card		12	4
Please improve the PMS agent forced installation / execution problem.		12	2
Shuttle bus time adjustment (+ extended weekend)		12	0
Activate bicycle road	*in-depth*,	11	3
I wish there was an air cleaner in every classroom.	*not feasible*,	11	2
Operate dormitory stores on Saturdays		10	1
Open both Korean and English lectures at the same time. .	*not feasible*, *creates discussion*,	9	0
Disabled-Friendly Campus	*original*, *in-depth*,	9	1
Eoeun-dong Mountain Underground Shopping Center . . .	*original*, *not feasible*,	9	4
Weekend dinner until 7:30	*not feasible*,	7	0
Sell more fruits in the store	*original*,	7	1
Small amounts of food at a small price!	*original*,	7	1
Let's have a party in Hoaam Dorm	*original*, *creates discussion*,	6	2
The library smells so bad.	*in-depth*,	6	0
Carillon bells are so mellow.	*original*,	6	0
Make a day of vegetarian meals once a week.	*original*, *in-depth*, *creates discussion*,	5	3
Daejeon Station → School Taxi App	*original*,	5	1
Solving the Outsiders Problem	*in-depth*, *creates discussion*,	4	2
Smoking map	*creates discussion*,	4	0
Bicycle roads on both sides of the roads	*original*,	4	1
Taxi stand platform congestion reminder	*original*,	4	0
Please cooperate with various online lecture sites	*in-depth*,	3	1
Please make a table tennis class	*not feasible*,	3	1
Access the building through a different type of identification when my student ID is gone.	*original*,	3	2
Please increase the amount of tuna kimbap in the shop. . .	*creates discussion*,	2	1
Electric Kickboard Registration Problem	*original*,	1	0
Installation so that dove can not sit on dormitory window frame	*original*, *creates discussion*,	0	0

Table 6.6: Significant ideas along with popularity (L = Likes, C = Comments) and rater's judgment, sorted by popularity.

7

Discussion

My results show some evidence for the applicability of the Self-Determination Theory (SDT) in the design of crowd-civic systems. However, there are still a number of open challenges regarding personalization. In this section, I discuss the results and limitations.

7.1 Benefits and Challenges of Theory-based Design

SDT has proven to be a useful perspective for designing applications dealing with voluntary participation. Learning from theory, as opposed to examples, can broaden the designer's view. I encourage other designers to consider human behavior theory to make more informed decisions about design of crowd-civic systems.

However, translating theory into concrete designs has been challenging. Theories validated in controlled laboratory environments may not perform as expected in an uncontrolled field experiment due to the interaction between a variety of factors. My attempt of translating SDT into distinct motivation-supportive messages was my best effort, but due to the subjective nature of motivation and behavior, it is not guaranteed that every design is perceived by the participants exactly as theorized. It is unlikely one can ever find the "perfect" solution here, and researchers have to rely on incremental improvements and continuous measurements.

Moreover, theory is often manifested in more than one concrete implementation. In this work, I used the Motivation to Volunteer Scale (MVS) and General Causality Orientations Scale (GCOS) questionnaires to capture two different constructs from the same theory. According to my results, both of these showed some correlations to both self-reported behavior (Study 1) and actual behavior (Study 2). While previous studies implied that both constructs should be applicable to voluntary work [14, 12], in Study 2, MVS has proven to be more useful to post-hoc classify users into two groups. It is possible that the MVS scale worked better here because its context is explicitly volunteering, whereas GCOS scale's context is mostly professional work. I had expected the GCOS' *Autonomy* and *Control* orientations to be closely related to a user's susceptibility of my two motivational treatments, but perhaps the tendency towards

certain regulatory styles is more indicative in the context of voluntary participation. It is also possible that this is a matter of the population that I studied: well educated university members might differ more in their volunteering motivations than in their general orientations.

7.2 Limitations and Alternatives for Message Personalization

According to the results from my studies, personalization of messages may be possible: treatment conditions showed significant differences in both studies, and personality data obtained through full questionnaires showed possible interaction effects between treatment and individual motivation. However, my attempt in Study 2 to personalize based on a subset of questions, backed by findings from Study 1, was unsuccessful. A likely explanation is that the classification using a subset of questions (4 out of 63) did not accurately capture individual motivations and orientations. Reducing the number of questions was a design trade-off I chose to reduce the risk of causing dropout or biasing users to think about their motivation. Successful personalization may depend on choosing the right method of data elicitation to feed the personalization algorithm. It is also possible that app users did not respond accurately to the personality test questions as they were taken out of a traditional survey context. The interface included an explanation to users that the questionnaire mattered, but I cannot be sure that that had the intended effect.

Further investigation of different ways of eliciting sufficient data at signup time to accurately classify users is needed. Alternatively, there may be ways to gather the data necessary to personalize without the need of the user's explicit input, for example by automatically inferring users' personalities from their social media profiles [13]. Furthermore, instead of one-shot personalization, one could use a dynamic system that learns the user's preferences while they are using the application. Similar attempts have been introduced in the context of online education [51]. This will require more careful and different motivational strategies, as one wouldn't be able to change affordances in the mutually exclusive way I did in this study while the user already started using the system. One might assign users to a motivation-neutral treatment at first, and only introduce motivation support once the system learned enough about the user to make a confident prediction.

However, automatic classification and personalization does not only have advantages. Users may be concerned about the involved privacy risks and scared by computers' increasing ability to judge and predict human behavior, and lament the lack of autonomy [49], something that self-determination theory clearly stipulates. Thus, personalization itself can be limiting, and its unintended side effects have to be investigated.

We may also consider a combination of personalization and customization by offering users a choice between different motivational affordances. This may mitigate the limitations of personality assessment while providing a form of autonomy to users. For example, a system might offer a choice between various affordances that support different kinds of motivations (similar to [19]). In turn, the user's choice could be used to infer their orientations and provide supportive messages in other parts of

the system. However, this can introduce risks of undermining motivation: users who were originally autonomously motivated might be negatively affected by being offered a choice of an extrinsically motivated regulation.

7.3 Studying Crowd-civic Systems

Researching design that affects behavior of voluntary crowds is challenging. In a controlled setting like Study 1, effects may be more apparent, but at the risk of lacking ecological validity. The reliability of self-reported intentions is also questionable due to the hypothetical bias [36]. In a natural environment like Study 2, one can expect greater validity, but measuring specific effects is harder due to a lack of control. By investigating my research questions in both settings, I attempted to provide a more balanced view. Research in crowd-civic systems is likely to benefit from both kinds of experiments.

Furthermore, the complexity and intertwining of motivation factors in voluntary work make empirical studies in this field difficult. Civic engagement is also a long-term effect affected by people's firmly held beliefs. I targeted a minimum of one week for the study to balance this to some extent, but running longitudinal studies will be beneficial.

I have gathered substantial evidence addressing my hypotheses, but additional experiments are needed to verify them. For example, survey data indicates some support for the theory that providing too strong external regulation can diminish autonomous motivation, which will lead to satisficing and higher dropout. A more long-term and larger scale experiment is needed to measure if people actually behave as they said.

In some ways, my research probe has succeeded in exemplifying a basic crowd-civic system. Users joined my platform with diverse motivations and they interacted with each other as a crowd to achieve civic goals. Both the MTurk and the campus community populations, though being biased each in their own way, showed a wide range of motivations. As the underlying theory of motivation is universal, I argue that the idea of motivation-supportive design is likely to generalize to other populations and other types of systems in this domain. However, the concrete implementation will vary and require separate testing, as motivation-support is not only dependent on the individual, but also on the specific task and overall goal.

My results make a contribution to the increasing calls for designing with diverse crowds in mind [20]. Civic platforms should support multiple motivations, not only in terms of different "factors" [2] (such as moral obligations, learning, fun, and recognition), but in terms of different regulatory styles relating to self-determination. We argue that this can increase not only overall participation, but also representativeness and inclusiveness. In this work, we showed how people with different motivations might be treated in a more supportive way. Future work could investigate how this affects representativeness by other factors, such as inclusiveness across demographics or variety of opinions.

8

Conclusion and Future Work

In this work, I investigated effects of motivation-supportive messages in crowd-civic systems. Informed by the Self-Determination theory, I created several messages for a number of tasks related to idea generation for social communities. In two studies, I gathered evidence for the applicability of this theory as well as a possibility of personalization. This work has implications for the design of systems that benefit from engagement from diverse groups. Especially in crowd-civic systems, this can increase not only overall participation, but also representativeness and inclusiveness.

Future work includes further iterations on message contents and presentation as well as more long-term studies. For the field study, even though the data exhibits promising trends, more participants are needed to draw definite conclusions. Regarding personalization, future work needs to investigate further the trade-offs between explicit and implicit data elicitation, personalization and customization, and affecting short- and long-term motivations. It will be interesting to see the effects of other interpretations of motivation support. For crowd-civic systems especially, versions lacking external rewards need to be investigated further. Neither is it always feasible nor desired to encourage participation with highly controlled regulation. Moreover, many effects might only become apparent in longer timeframes. Thus, running longitudinal studies would be beneficial. However, setting these up is challenging due to the complexity of this domain's design space.

To help future researchers to build on top of my work in this domain, I open-sourced the application developed for this study. The source code for all parts of this work (survey, web application frontend and backend, and R analyses) can be found on GitHub: http://github.com/graup/manyideas/. My hope is that similar studies can be run in other environments and the notion of motivation-supportive personalization be iteratively improved.

Appendices

Appendix A Survey questions

A.1 Study 1

0. **Welcome**
    1.) Sex (Female, Male)
    2.) Age (18-26, 27-32, 33-40, 41-55, 56+)
    3.) Highest attained level of education (No high school, High school / GED, Some college, Bachelor's degree, Master's degree, PhD or higher)

1. **GCOS**[1] (All items require responses on a scale of 1-7, *very unlikely* to *very likely*. Items were shown in this order. Answer scale key here shown in braces.)
    1.) You are taking a personality test. You are likely to:
        i. Read all the items thoroughly
        ii. Pay attention and think about each answer before choosing a response
        iii. Rate answers randomly because you don't care
    2.) You have been offered a new position in a company where you have worked for some time. The first question that is likely to come to mind is:
        i. What if I can't live up to the new responsibility? (I)
        ii. Will I make more at this position? (C)
        iii. I wonder if the new work will be interesting (A)
    3.) You have a school-age daughter. On parents' night the teacher tells you that your daughter is doing poorly and doesn't seem involved in the work. You are likely to:
        i. Talk it over with your daughter to understand further what the problem is. (A)
        ii. Scold her and hope she does better. (I)
        iii. Make sure she does the assignments, because she should be working harder. (C)
    4.) You had a job interview several weeks ago. In the mail you received a form letter which states that the position has been filled. It is likely that you might think:
        i. It's not what you know, but who you know. (C)
        ii. I'm probably not good enough for the job. (I)
        iii. Somehow they didn't see my qualifications as matching their needs. (A)
    5.) You are a plant supervisor and have been charged with the task of allotting coffee breaks to three workers who cannot all break at once. You would likely handle this by:
        i. Telling the three workers the situation and having them work with you on the schedule. (A)
        ii. Simply assigning times that each can break to avoid any problems. (C)
        iii. Find out from someone in authority what to do or do what was done in the past. (I)
    6.) A close (same-sex) friend of yours has been moody lately, and a couple of times has become very angry with you over "nothing." You might:
        i. Share your observations with him/her and try to find out what is going on for him/her. (A)
        ii. Ignore it because there's not much you can do about it anyway. (I)
        iii. Tell him/her that you're willing to spend time together if and only if he/she makes more effort to control him/herself. (C)
    7.) You have just received the results of a test you took, and you discovered that you did very poorly. Your initial reaction is likely to be:
        i. "I can't do anything right," and feel sad. (I)
        ii. "I wonder how it is I did so poorly," and feel disappointed. (A)
        iii. "That stupid test doesn't show anything," and feel angry. (C)
    8.) You have been invited to a large party where you know very few people. As you look forward to the evening, you would likely expect that:
        i. You'll try to fit in with whatever is happening in order to have a good time and not look bad. (C)
        ii. You'll find some people with whom you can relate. (A)
        iii. You'll probably feel somewhat isolated and unnoticed. (I)

---

[1]   http://selfdeterminationtheory.org/general-causality-orientations-scale/

9.) You are asked to plan a picnic for yourself and your fellow employees. Your style for approaching this project could most likely be characterized as:

    i. Take charge: that is, you would make most of the major decisions yourself. (C)

    ii. Follow precedent: you're not really up to the task so you'd do it the way it's been done before. (I)

    iii. Seek participation: get inputs from others who want to make them before you make the final plans. (A)

10.) Recently a position opened up at your place of work that could have meant a promotion for you. However, a person you work with was offered the job rather than you. In evaluating the situation, you're likely to think:

    i. You didn't really expect the job; you frequently get passed over. (I)

    ii. The other person probably "did the right things" politically to get the job. (C)

    iii. You would probably take a look at factors in your own performance that led you to be passed over. (A)

11.) For an online study, you have to read and answer a set of questions. You are likely to:

    i. Read every scenario and think about it thoroughly

    ii. Answer everything as quickly as possible without thinking much

    iii. Think about each answer and rate them honestly

12.) You are embarking on a new career. The most important consideration is likely to be:

    i. Whether you can do the work without getting in over your head. (I)

    ii. How interested you are in that kind of work. (A)

    iii. Whether there are good possibilities for advancement. (C)

13.) A woman who works for you has generally done an adequate job. However, for the past two weeks her work has not been up to par and she appears to be less actively interested in her work. Your reaction is likely to be:

    i. Tell her that her work is below what is expected and that she should start working harder. (C)

    ii. Ask her about the problem and let her know you are available to help work it out. (A)

    iii. It's hard to know what to do to get her straightened out. (I)

14.) Your company has promoted you to a position in a city far from your present location. As you think about the move you would probably:

    i. Feel interested in the new challenge and a little nervous at the same time. (A)

    ii. Feel excited about the higher status and salary that is involved. (C)

    iii. Feel stressed and anxious about the upcoming changes. (I)

2. **MVS** [14] (All items require responses on a scale of 1-5, *corresponds not at all* to *corresponds to a great extent*. Items were shown in randomized (per user) order. Answer scale key here shown in braces.)

1.) for the pleasure I feel in finding new ways of help (intrinsic)

2.) for the pleasure and interest I feel in doing this activity (intrinsic)

3.) for the pleasure I feel in doing something new (intrinsic)

4.) for the pleasure I feel when I master the situations I'm dealing with (intrinsic)

5.) because this activity has become an integral part of my life (integrated)

6.) because volunteering has become a part of who I am (integrated)

7.) because it is one of the ways I live my life (integrated)

8.) because volunteering is a suitable activity for me (integrated)

9.) because it's something that is fulfilling for me as a person (identified)

10.) because it's something that contributes to my personal growth (identified)

11.) because it is a wise thing to do (identified)

12.) because it's a good way to contribute (identified)

13.) because I would feel very bad if I did not help others (introjected)

14.) because I would feel guilty if I did not volunteer (introjected)

15.) because I would regret not doing volunteering (introjected)

16.) because I would be ashamed if I did not volunteer (introjected)

17.) because other people will be sorry if I didn't do it (external)

18.) for the recognition I get from others (external)

19.) to avoid being criticized (external)

20.) because I know others are pleased that I volunteer (external)

21.) I don't know; Sometimes I have the impression I'm wasting time when I volunteer (amotivation)

22.) I don't know; I can't see how my efforts are helping others when I volunteer (amotivation)

23.) I don't know; I can't see how all this helps (amotivation)

24.) I don't know; I can't see what I'm getting out of it (amotivation)

3. **Comparison** (22 pairwise comparisons. $\binom{7}{2} + 1 = 22$ A/B preferences covering all combinations and one repeated comparison as an attention check, seeing only one combination at a time. Order randomized per user using balanced latin square method.)

4. **Post-comparison**
    1.) According to your ratings, this is the version that you prefer the most. (Screenshot)
       Is this really a version you liked? (Yes, No)
    2.) Please explain in one sentence why you think you preferred (or did not prefer) this version. (Free-form text)
    3.) Thinking more generally, what would motivate you personally to participate in this app? Feel free to share any ideas or thoughts you have. (Free-form text)

A.2 Study 2

0. **Welcome**
    1.) Email address used to sign up in app
1. **GCOS** (see Study 1 survey)
2. **MVS** (see Study 1 survey)
3. **App Usage** (If not otherwise noted, scale 1-7, *corresponds not all* to *corresponds completely*. Item short names shown in braces.)
    1.) What motivated you to join this app? (reason_join)
          i. I thought the app's goal sounded interesting. (interesting)
         ii. I thought using the app could be fun. (fun)
        iii. I thought the app's goal is important. (important)
         iv. I wanted to see what other people posted. (relatedness)
          v. I wanted to share my own idea. (sharing)
         vi. I wanted to help this research project. (help)
        vii. I was interested in the offered rewards. (reward)
       viii. I don't really know why I joined. (amotivation)
    2.) Any other reason? (Free-form text)
    3.) After your first session, did you plan on revisiting the app? (Unlikely 1 - 5 Very likely)
    4.) The incentive period has finished, but would you be interested in continuing to use this application? (Unlikely 1 - 5 Very likely)
    5.) Without any special events, how often would you check the app? (Never 1 - 5 Daily)
    6.) If you stopped using the app, why? (reason_dropout)
          i. The content wasn't interesting enough. (interesting)
         ii. I didn't know what to do. (unclear)
        iii. The rewards offered were too little. (reward)
         iv. It was hard to imagine what my contributions will be useful for. (impact)
          v. I didn't have more time. (time)
         vi. I thought I had contributed enough. (enough)
    7.) If you could choose your own reward, what would that be? (Free-form text)
    8.) Apart from a reward, what else would motivate you to continue using this app? (Free-form text)
4. **AttrakDiff Mini** [15]
    1.) With the help of these word pairs, please enter what you consider the most appropriate description for Many Ideas. (Scale 1-7 with the two words shown at the two ends of each scale. Scale names here shown in braces. Items were shown in this order, but word pairs and scales were randomly swapped in polarity.)
          i. complicated — simple (pq)
         ii. ugly — attractive (beauty)
        iii. impractical — practical (pq)
         iv. tacky — stylish (hq)
          v. unpredictable — predictable (pq)
         vi. cheap — premium (hq)
        vii. unimaginative — creative (hq)
       viii. bad — good (goodness)
         ix. confusing — clearly structured (pq)
          x. dull — captivating (hq)
    2.) Do you have any suggestions how we could make this app better? (Free-form text)

Appendix B Design Artifacts

B.1 Online Survey Layout

Figure 1: Layout of start page of survey.

Figure 2: Layout of General Causality Orientation Score questionnaire, step one of survey.

Completed 45 of 93 questions (48%)

Step 2 / 4

Great job! We appreciate your honest answers.

The following questions are about volunteering. Please try to answer them as good as you can.

In the past three months, how often, on average, did you participate in volunteering activities?

Never ▲▼

Please indicate to what extent each of the following items correspond to your personal motives for engaging in volunteering.

I volunteer...

1. because it's a good way to contribute

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

2. because it is a wise thing to do

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

3. I don't know; I can't see how all this helps

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

4. because I would regret not doing volunteering

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

24. because I would feel very bad if I did not help others

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

25. To show that you are concentrated, please answer this question with 5.

○	○	○	○	○
1	2	3	4	5

corresponds   corresponds
not at all   to a great extent

Continue

Figure 3: Layout of Motivation to Volunteer Scale questionnaire, step two of survey.

Figure 4: Layout of instructions for comparison task, step three of survey.

Figure 5: Layout of comparison task, step three of survey.
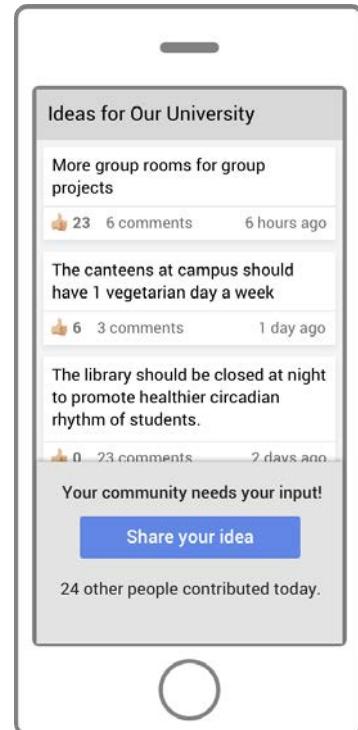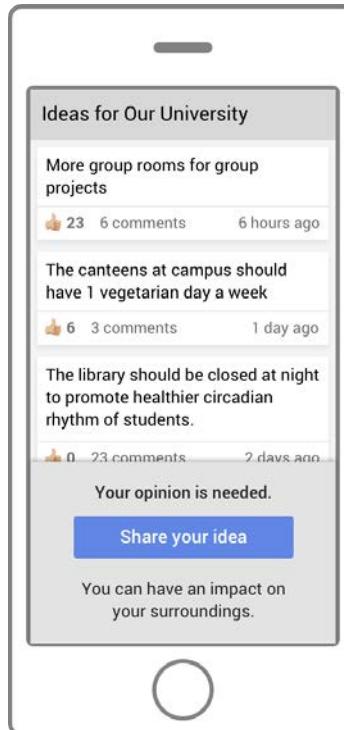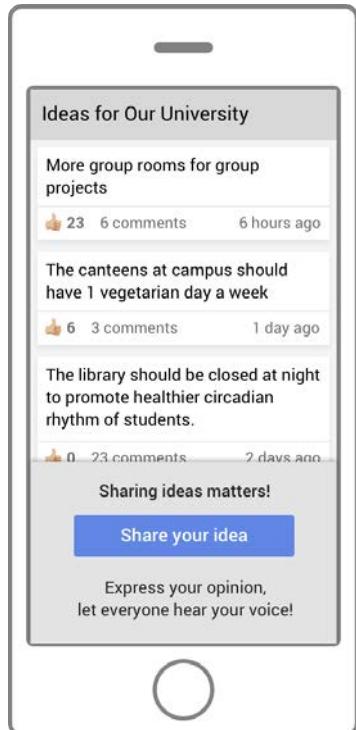
Figure 6: Layout of post-comparison questions, step four of survey.

B.2 Online Survey Comparison Versions

(a) Need for autonomy supportive message.

(b) Need for competence supportive message

(c) Need for relatedness supportive message

(d) Autonomous orientation supportive message

(e) Impersonal orientation supportive message

(f) Controlled orientation supportive message.

Figure 7: All mockup versions for Study 1 pairwise comparison task.

B.3 Many Ideas Motivation-supportive Messages

Table 1: Case study app components with motivation-supportive message treatment versions, in the order that users are likely to perceive them.

Baseline	Autonomy-supportive	Control-supportive
(Main feed) Call-to-Action		
[Share your idea] We need everyone's contribution!	[Share your idea] <N> other people contributed today. Your ideas matter.	Become a Contributor of the Week! [Share your idea] Your chance to win $20!
(New idea) Intro text		
(No text)	Our voice matters! Among all ideas submitted until April 11, we will hand-pick three promising ideas, present them to the whole school, and follow up with concrete steps to support their implementation! Together, we can have a real impact.	Among all contributors until April 11th, we will randomly select 10 members to win $20.
(Main feed) Tutorial Message		
Welcome to Many Ideas for [Uni]! We're happy to have you here. On this page, you can see ideas by other members. Do you see anything you are interested in? Try tapping on a post.	Welcome to Many Ideas for [Uni], **where our combined voice can have a real impact**! We're happy to have you here. On this page, you can see ideas by other members. Do you see anything you are interested in? Try tapping on a post.	Welcome to Many Ideas for [Uni]! We're happy to have you here. On this page, you can see ideas by other members. Do you see anything you are interested in? Try tapping on a post. **Participating actively can get you a chance to win $20!**
(View others' idea) Tutorial Message		
Here you can see details about the idea and comments that other people left. If you agree with this issue, how about showing your support by tapping the heart? If you have another opinion, try writing a short comment.	Here you can see details about the idea. **Discussing ideas and showing your support is important for our community.** If you agree with this issue, how about showing your support by tapping the heart? If you have another opinion, try writing a short comment.	Here you can see details about the idea. If you agree with this issue, how about showing your support by tapping the heart? If you have another opinion, try writing a short comment. **Remember, all active users have a chance to win $20!**
(My ideas) Empty State		
Nothing yet. Post your first idea now!	Nothing yet. Share your ideas with the community now! Everyone's voice counts.	Nothing yet. Share your ideas for a chance to win $20!
(FAQ) Why should I participate? section		
You can have an impact on improving conditions for everyone. We are also preparing several rewards for good ideas and active contributors. We will share the details with you later, but rest assured all active participants are always eligible for all rewards.	You can have a real impact on improving conditions for everybody! **[Reward description text similar to corresponding Intro text]** As part of our study, we also consider other rewards. We cannot show you that right now, but all active participants will always be eligible for all rewards.	

Tutorial Messages are only shown until the prompted action has been performed by the user for the first time.
Bold text in this table was also highlighted similarly in the application to make the treatment messages more salient.

List of Figures

List of Tables

Bibliography

[1] Tanja Aitamurto and Hélène Landemore. Crowdsourced deliberation: The case of the law on off-road traffic in finland. *Policy & Internet*, 8(2):174–196, 2016.

[2] Tanja Aitamurto and Jorge Saldivar. Motivating participation in crowdsourced policymaking: The interplay of epistemic and interactive aspects. In *Proceedings of the 21st ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 2017.

[3] Chris Bateman. Playing work, or gamification as stultification. *Information, Communication & Society*, pages 1–11, 2018.

[4] Kirsten Boehner and Carl DiSalvo. Data, design and civics: an exploratory study of civic tech. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2970–2981. ACM, 2016.

[5] Diana I Cordova and Mark R Lepper. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of educational psychology*, 88(4): 715, 1996.

[6] Lorrie Faith Cranor. I didn't buy it for myself. In *Designing personalized user experiences in eCommerce*, pages 57–73. Springer, 2004.

[7] Herbert Aron David. The method of paired comparisons. *Fifth Conference on the Design of Experiments in Army Research Development and Testing*, 1959.

[8] Edward L Deci and Richard M Ryan. The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, 19(2):109–134, 1985.

[9] Menno Deen. *GAME Games Autonomy Motivation & Education: How autonomy-supportive game design may improve motivation to learn*. PhD thesis, Technische Universiteit Eindhoven, Eindhoven, NL, 2015.

[10] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM, 2011.

[11] Sheena Erete and Jennifer O Burrell. Empowered participation: How citizens use technology in local governance. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2307–2319. ACM, 2017.

[12] Marylène Gagné and Edward L Deci. Self-determination theory and work motivation. *Journal of Organizational behavior*, 26(4):331–362, 2005.

[13] Jennifer Golbeck, Cristina Robles, and Karen Turner. Predicting personality with social media. In *CHI'11 extended abstracts on human factors in computing systems*, pages 253–262. ACM, 2011.

[14] Caterina Grano, Fabio Lucidi, Arnaldo Zelli, and Cristiano Violani. Motives and determinants of volunteering in older adults: An integrated model. *The International Journal of Aging and Human Development*, 67(4):305–326, 2008.

[15] Marc Hassenzahl and Andrew Monk. The inference of perceived usability from beauty. *Human–Computer Interaction*, 25(3):235–260, 2010.

[16] Reinhold Hatzinger, Regina Dittrich, et al. Prefmod: An r package for modeling preferences based on paired comparisons, rankings, or ratings. *Journal of Statistical Software*, 48(10):1–31, 2012.

[17] Natali Helberger, Kari Karppinen, and Lucia D'Acunto. Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2):191–207, 2018.

[18] Gary Hsieh and Rafał Kocielnik. You get who you pay for: The impact of incentives on participation bias. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 823–835. ACM, 2016.

[19] Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voida. Personality-targeted gamification: a survey study on personality traits and motivational affordances. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2001–2013. ACM, 2016.

[20] Eunice Jun, Gary Hsieh, and Katharina Reinecke. Types of motivation affect study selection, attention, and dropouts in online experiments. In *Proceedings of the 21st ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 2017.

[21] Clare-Marie Karat, Jan O Blom, and John Karat. *Designing personalized user experiences in eCommerce*, volume 5. Springer Science & Business Media, 2004.

[22] John Karat, Clare-Marie Karat, and Carolyn Brodie. Personalizing interaction. In *Designing personalized user experiences in eCommerce*, pages 7–17. Springer, 2004.

[23] Juho Kim. *Learnersourcing: improving learning with collective learner activity*. PhD thesis, Massachusetts Institute of Technology, 2015.

[24] Nam Wook Kim, Jonghyuk Jung, Eun-Young Ko, Songyi Han, Chang Won Lee, Juho Kim, and Jihee Kim. Budgetmap: Engaging taxpayers in the issue-driven classification of a government budget. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1028–1039. ACM, 2016.

[25] Rafal Kocielnik and Gary Hsieh. Send me a different message: Utilizing cognitive space to create engaging message triggers. In *CSCW*, pages 2193–2207, 2017.

[26] Richard Koestner, Gaëtan F Losier, Robert J Vallerand, and Donald Carducci. Identified and introjected forms of political internalization: Extending self-determination theory. *Journal of personality and social psychology*, 70(5):1025, 1996.

[27] Sherrie YX Komiak and Izak Benbasat. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS quarterly*, pages 941–960, 2006.

[28] Markus Krause, Marc Mogalle, Henning Pohl, and Joseph Jay Williams. A playful game changer: Fostering student retention in online education with social gamification. In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale*, pages 95–102. ACM, 2015.

[29] Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. *Building successful online communities: Evidence-based social design*. Mit Press, 2012.

[30] Vinodh Krishnaraju, Saji K Mathew, and Vijayan Sugumaran. Web personalization for user acceptance of technology: An empirical investigation of e-government services. *Information Systems Frontiers*, 18(3):579–595, 2016.

[31] David Martin, Sheelagh Carpendale, Neha Gupta, Tobias Hoßfeld, Babak Naderi, Judith Redi, Ernestasia Siahaan, and Ina Wechsung. Understanding the crowd: Ethical and practical matters in the academic use of crowdsourcing. In *Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments*, pages 27–69. Springer, 2017.

[32] Brian McInnis, Alissa Centivany, Juho Kim, Marta Pobet, Karen Levy, and Gilly Leshed. Crowdsourcing law and policy: A design-thinking approach to crowd-civic systems. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW'17 Companion)*, page 355, 2017.

[33] 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.

[34] Giovanna Menardi and Nicola Torelli. Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery*, 28(1):92–122, 2014.

[35] Youngme Moon. Personalization and personality: Some effects of customizing message style based on consumer personality. *Journal of Consumer Psychology*, 12(4):313–325, 2002.

[36] James J Murphy, P Geoffrey Allen, Thomas H Stevens, and Darryl Weatherhead. A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3): 313–325, 2005.

[37] BPS Murthi and Sumit Sarkar. The role of the management sciences in research on personalization. *Management Science*, 49(10):1344–1362, 2003.

[38] Babak Naderi. *Motivation of Workers on Microtask Crowdsourcing Platforms*. Springer, 2018.

[39] Oded Nov and Ofer Arazy. Personality-targeted design: Theory, experimental procedure, and preliminary results. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 977–984. ACM, 2013.

[40] Ory Okolloh. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory learning and action*, 59(1):65–70, 2009.

[41] Maria V Palacin-Silva, Antti Knutas, Maria Angela Ferrario, Jari Porras, Jouni Ikonen, and Chandara Chea. The role of gamification in participatory environmental sensing: A study in the wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 221. ACM, 2018.

[42] Mayur Patel, Jon Sotsky, Sean Gourley, and Daniel Houghton. Knight foundation: Trends in civic tech, 2014. URL http://www.knightfoundation.org/features/civictech. Accessed: 2018-01-22.

[43] Anne C Rouse. A preliminary taxonomy of crowdsourcing. *ACIS 2010 Proceedings*, 76:1–10, 2010.

[44] Richard M Ryan and Edward L Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1):54–67, 2000.

[45] Amir Salihefendic. How reddit ranking algorithms work, 2015. URL https://medium.com/hacking-and-gonzo/how-reddit-ranking-algorithms-work-ef111e33d0d9. Accessed: 2018-05-23.

[46] Saiph Savage, Andres Monroy-Hernandez, and Tobias Höllerer. Botivist: Calling volunteers to action using online bots. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 813–822. ACM, 2016.

[47] Anne M Sinatra. A review of self-reference and context personalization in different computer-based educational domains. *Design Recommendations for Intelligent Tutoring Systems*, page 107, 2016.

[48] Kar Yan Tam and Shuk Ying Ho. Understanding the impact of web personalization on user information processing and decision outcomes. *MIS quarterly*, pages 865–890, 2006.

[49] Eran Toch, Yang Wang, and Lorrie Faith. Personalization and privacy : a survey of privacy risks and remedies in personalization-based systems. pages 203–220, 2012. doi: 10.1007/s11257-011-9110-z.

[50] Lars Hasselblad Torres. Citizen sourcing in the public interest. *Knowledge Management for Development Journal*, 3(1):134–145, 2007.

[51] Joseph Williams, Na Li, Juho Kim, Jacob Whitehill, Samuel Maldonado, Mykola Pechenizkiy, Larry Chu, and Neil Heffernan. The mooclet framework: Improving online education through experimentation and personalization of modules. 2014.

[52] Ping Zhang. Technical opinion motivational affordances: reasons for ict design and use. *Communications of the ACM*, 51(11):145–147, 2008.

[53] Dieter Zinnbauer. Crowdsourced corruption reporting: What petrified forests, street music, bath towels, and the taxman can tell us about the prospects for its future. *Policy & Internet*, 7(1):1–24, 2015.