Doctoral Thesis

Human Behavior Experimentation and Participation in Scientific Activities in the Wild

by Julián A. Vicens Bennasar

UNIVERSITAT ROVIRA i VIRGILI

Doctoral Thesis

Human Behavior Experimentation and Participation in Scientific Activities in the Wild

Julián A. Vicens Bennasar

Supervised by

Dra. Mercè Gisbert Cervera

Dr. Jordi Duch Gavaldà

Departament de Pedagogia

March 2018

Tarragona

UNIVERSITAT ROVIRA i VIRGILI

I STATE that the present thesis, entitled **"Human Behavior Experimentation and Participation in Scientific Activities in the Wild"** and presented by **Julián A. Vicens Bennasar** for the award of the degree of Doctor, has been carried out under my supervision at the *Departament de Pedagogia de la Universitat Rovira i Virgili* and meet the requirements to opt to International Mention.

Tarragona.
March 1, 2018

The doctoral thesis supervisors:

Dra. Mercè Gisbert Cervera

Dr. Jordi Duch Gavaldà

Acknowledgments

M'agradaria agrair a sa gent que durant aquests quatre anys de treball m'heu ajudat d'una manera o una altra a dur a port aquesta tesi.

Als meus directors, a na Mercè Gisbert per fer-me un lloc al teu laboratori, per sa teva hospitalitat a Tarragona des d'es primer moment i per sempre cercar sa part educativa de qualsevol cosa que faci. I a en Jordi Duch per agafar es timó i guiar-me, per introduir-me a noves experiències al món de la recerca. Però sobretot pel teu tracte humà i proper.

A n'Imma Fontdevila i a n'Artur Duart, que vau fer un espai per jo a "ca vostra" i amb qui he compartit llargues converses sobre tecnologia, cultura i més. Gràcies per tenir, periòdicament, una estona per jo.

A tots i cada un dels companys del Late, especialment a na Vane Esteve, en Francesc Esteve, n'Anna Sanchez, na Mònica Sanromà i en Jose Cela, amb qui he compartit grans estones i moltes rialles.

Gràcies Josep Perelló per obrir-me ses portes d'OpenSystems, i per compartir sa teva visió de sa ciència més enllà de sa ciència. Muchas gracias Anna Cigarini por estar siempre tan comprometida y hacer muy fácil trabajar contigo.

Gracias Anxo Sanchez por tu inagotable disponibilidad para ayudar, por tu amabilidad y por tu hospitalidad en el GISC. Gracias también a Ignacio Tamarit, Alberto Antonioni y Maria Pereda.

Als companys de "complex systems" de Alephys i UBICS que, sense sabre-ho, m'heu ensenyat una manera alternativa de veure el mon.

I would like to thank Haoqi Zhang, the co-directors and colleagues of Delta Lab for welcoming me. Special thanks to Daniel Rees Lewis (for finally having that beer), Yongsung Kim and Eureka Foong. I would also like to thank the people of AmaralLab for the inspiring discussions. Muchas gracias Julia Poncela y Helio Tejedor por vuestra hospitalidad y generosidad en Chicago.

Als companys de Magma, especialment a na Mariela Balbona, na Susana Bernadi, en Jaume Benabarre, n'Anna Ruiz, na Sandra Niubo i en Carles Gisbert, tot era molt més lleuger i divertit amb voltros.

Y además a toda la gente con la que he colaborado Yamir Moreno, Carlos Gracia, Jesus Gómez, Nereida Bueno, he aprendido mucho de vosotros. Igualment amb na Mar Camacho, Isabelle Bonhoure, Enric Senabre i Mario Gutiérrez.

I sobretot a totes ses persones que han participat a ses experiències de ciència ciutadana i que són font elemental d'aquesta recerca.

A na Míriam pels teus consells lingüístics (i no lingüístics).

Gràcies a sa meva família, al·lota i amics, què també sou família.

Julian A. Vicens Bennasar
March 1, 2018

List of Contributions

1. Poncela-Casasnovas, J., Gutiérrez-Roig, M., Gracia-Lázaro, C., Vicens, J., Gómez-Gardeñes, J., Perelló, J., Moreno, Y., Duch, J., Sánchez, A. "Humans Display a Reduced Set of Consistent Behavioral Phenotypes in Dyadic Games." *Science Advances*, **2**(8). (2016).

2. Cigarini A., Vicens J., Duch J., Sánchez A, Perelló J. Quantitative Account of Social Interactions in a Mental Health Care Ecosystem: Cooperation, Trust and Collective Action." *Scientific Reports*, **1**(8). (2018).

3. Vicens, J., Bueno-Guerra, N., Gutiérrez-Roig, M., Gracia-Lázaro, C., Gómez-Gardeñes, J., Perelló, J., Sánchez, A., Moreno, Y., Duch, J. "Resource Heterogeneity Leads to Unjust Effort Distribution in Climate Change Mitigation." *PLoS ONE*, **13**(10):e0204369. (2018).

4. Vicens, J., Zhang, H., Duch, J. "Natural Patterns: A Platform To Enhance Participation And Science Disposition In Citizen Science Activities." (2018). (Submitted)

5. Vicens, J., Perelló, J., Duch, J. "Citizen Social Lab: A Digital Platform for Human Behaviour Experimentation Within a Citizen Science Framework." (2018). (Submitted)

. Last update November 1th, 2018

Abstract

Cooperation is one of the behavioral traits that define human beings – and other complex systems – and that has allowed us to evolve. However, even nowadays, after years of scientific advances we are still trying to understand why some systems, and particularly humans, cooperate. Human behavior lab experiments based on social dilemmas modeled as behavioral games try to shed light on these unknowns.

One of the recent approaches in this vein has been the move of human behavior experimentation from laboratories to public spaces, where behaviors occur naturally, to study the main behavioral traits – cooperation, trust, reciprocity, risk aversion or collectivity sense –. Citizen science practices have provided a perfect framework to experiment in the wild and to promote the participation of people that are usually not involved in science.

This dissertation is focused on advancing the field of behavioral experimentation in the open environments by performing experiments based on citizen science practices. The work is divided in two blocs, one oriented to the design of experimental platforms and another focused on the analysis of experimental results. Particularly in the first, we study how to design citizen science platforms that allow scientific activities in the wild and we present two of them that are applied to two different contexts. In the second, we perform experimental studies of social systems in order to analyze behavioural traits, to look for the emergency of behavioral patterns and also to evaluate the designs of the platforms.

With the first platform we investigate how citizen science systems can serve as a catalyst for the promotion of scientific thinking and to increase engagement in science. We introduce Natural Patterns, a platform that presents scientific activities based on the scientific research method. We introduce the results of a user experience study and, in light of the contributions of the participants, we describe a series of design principles that increase the motivation of the participants and promote scientific participation in the wild in this specific context.

The second platform is designed to study traits of human behavior and to help creating experiments in the wild, where real social phenomena emerge, by encouraging the participation within the framework of citizen science. To that effect, the platform is very modular and includes a broad range suite of behavior games. It is light, which allows experimentation in places with little infrastructure – public spaces, festivals or conferences –, and scalable, allowing to introduce new social dilemmas and interactions. We evaluate participation, robustness and quality of the collected data in all the experiments that have been carried out so far.

With respect to the experimental studies, we present the first two human behavior experiments that were carried out with the platform. In the first, participants have to predict market movements under different circumstances. Specifically, we study whether the strategies that emerge are robust in the replicas of the experiments that were performed in different locations with very different sociodemographic samples. The second experiment studies the behavior patterns – or phenotypes – that emerge from the decisions that participants make when they face a set of social dilemmas. Concretely, we focus on the data analysis by unsupervised learning techniques and conclude that the behavior of the participants facing this set of social dilemmas with different tensions can be summarized in a total of five phenotypes.

The last two experiments presented in this dissertation also tackle and fully delve into human behavior experimentation. The first is a collective-risk dilemma in which a group of participants has to contribute their assets to overcome a collective damage, specifically climate change. We study how resource inequalities cause unfair behaviors, provoking the most vulnerable participants to be more harmed than the most favored ones.

Finally, the last experiment is a collective-risk dilemma, but in this case there are no economic inequalities between the participants. Unlike the rest of experiments, this one is carried out within a given collective: the mental health ecosystem, formed mainly by people affected by mental disorders, caregivers and relatives. We study the tensions that exist when all of them collaborate collectively to solve a collective uncertain bad. From the results of the experiment it is emphasized that the cost of collective actions falls on those affected by mental conditions.

To conclude, the platforms that we have built help us to carry out participatory scientific activities within the framework of citizen science, which has plenty of applications, from learning to activism. This is especially so in the case of the experimentation platform that studies human behavior in vivo, which makes it possible to access a representative sample of the population and to integrate the participants in the research process. The platform will help set the basis for future behavioral experiments in the wild. In the light of the social experiments' results we can comprehend how we behave collectively, in society or within a certain group, when we face social dilemmas, and consequently evaluate behavioral traits and the emergence of behavioral patterns thanks to unsupervised machine learning techniques. These insights offer new ways to study the tensions that exist in a given community around a particular social issue.

Resum

La cooperació és un dels trets de comportament que defineixen els éssers humans, així com altres sistemes complexos, i que ens ha permès evolucionar com a espècie. No obstant això, fins i tot avui dia, després de segles d'avenços científics, encara tractem d'entendre per què cooperen alguns sistemes, i particularment per què ho fan els humans. Els experiments de comportament humà, que es duen a terme en laboratoris i es basen en dilemes socials modelats com a jocs de comportament, tracten de donar llum sobre aquests aspectes encara desconeguts.

Una de les aproximacions recents ha estat el trànsit de l'experimentació en comportament humà des de laboratoris cap a espais públics, on els comportaments ocorren de manera natural, per així estudiar els principals trets de comportament: cooperació, confiança, reciprocitat, aversió al risc o sentit col·lectiu. Les pràctiques de ciència ciutadana han proporcionat un marc perfecte per experimentar en el camp i promoure la participació de persones que generalment no estan involucrades en la pràctica científica.

Aquesta dissertació es centra en avançar en el camp de l'experimentació conductual en entorns oberts utilitzant experiments emmarcats en les pràctiques de ciència ciutadana. Aquest treball es divideix en dos blocs, un orientat al disseny de plataformes experimentals i un altre centrat en l'anàlisi de resultats experimentals. En el primer, concretament, estudiem com dissenyar plataformes de ciència ciutadana que permeten realitzar activitats científiques en entorns oberts i presentem dues d'elles que s'apliquen en dos contextos diferents. En el segon bloc, vam realitzar estudis experimentals de sistemes socials per analitzar trets del comportament, buscar l'emergència de patrons de comportament i en última instància per avaluar els dissenys de les plataformes.

Amb la primera plataforma investiguem com els sistemes de ciència ciutadana poden servir com a catalitzador per a la promoció del pensament científic i per augmentar la participació en activitats científiques. Presentem Natural Patterns, una plataforma que presenta activitats científiques basades en el mètode d'investigació científica. Presentem els resultats d'un estudi d'experiència de l'usuari i, a la llum de les contribucions dels participants, descrivim una sèrie de principis de disseny que augmenten la motivació dels participants i promouen la participació en activitats científiques a l'entorn natural.

La segona plataforma està dissenyada per estudiar trets del comportament humà i ajudar a crear experiments en entorns oberts, on sorgeixen fenòmens socials reals, i ho fem fomentant la participació en el marc de la ciència ciutadana. A aquest efecte, la plataforma és modular i inclou una àmplia gamma de jocs de comportament. És lleugera, el que permet l'experimentació en llocs amb poca infraestructura (espais públics, festivals o conferències) i escalable, la qual cosa permet introduir nous dilemes i interaccions socials. Avaluem la participació, la solidesa i la qualitat de les dades recopilades en tots els experiments que s'han dut a terme fins al moment.

Pel que fa als estudis experimentals, presentem els primers dos experiments de comportament humà que es van dur a terme amb la plataforma. En el primer, els participants han de predir els moviments del mercat financer en diferents circumstàncies. Específicament, estudiem si les estratègies que emergeixen són sòlides en els experiments rèplica que s'han realitzat en diferents llocs amb mostres sociodemogràfiques molt diferents. El segon experiment estudia els patrons de comportament (o fenotips) que sorgeixen de les decisions que prenen els participants quan s'enfronten a un conjunt de dilemes socials. Concretament, ens centrem en l'anàlisi de dades mitjançant tècniques de "unsupervised learning", del qual vam concloure que el comportament dels participants que prenen decisions en dilemes socials amb diferents tensions es pot caracteritzar en un total de cinc fenotips.

Els dos últims experiments presentats en aquesta dissertació també aborden i aprofundeixen en l'experimentació del comportament humà. El primer és un "collective-risk dilemma" en el qual un grup de participants ha de contribuir amb els seus actius per superar un dany col·lectiu, concretament el canvi climàtic. Estudiem com les desigualtats de recursos causen comportaments injustos, provocant que els participants més vulnerables es vegin més perjudicats que els més afavorits.

Finalment, l'últim experiment és un "collective-risk dilemma", però en aquest cas no hi ha desigualtats econòmiques entre els participants. A diferència de la resta d'experiments, aquest es va dur a terme dins d'un col·lectiu particular: l'ecosistema de salut mental, format principalment per persones afectades per trastorns mentals, cuidadors i familiars. Estudiem les tensions que existeixen quan totes elles col·laboren col·lectivament per resoldre un dany col·lectiu incert. A partir dels resultats de l'experiment, s'infereix que el cost de les accions col·lectives recau sobre els afectats per trastorns mentals.

Per concloure, les plataformes que hem desenvolupat ens ajuden a dur a terme activitats científiques participatives en el marc de la ciència ciutadana, que té aplicacions des de l'aprenentatge fins l'activisme. Això és especialment notori en el cas de la plataforma d'experimentació que estudia el comportament humà in vivo, i que fa possible accedir a una mostra representativa de la població i a integrar els participants en el procés d'investigació. La plataforma ajudarà a establir les bases per a futurs experiments de comportament en entorns oberts. A la llum dels resultats dels experiments socials podem comprendre com ens comportem col·lectivament, en societat o dins d'un determinat grup, quan enfrontem dilemes socials, i conseqüentment avaluem els trets de comportament i l'aparició de patrons de comportament gràcies a tècniques d'"unsupervised learning". Aquestes idees ofereixen noves formes d'estudiar les tensions que existeixen en una comunitat concreta al voltant d'un problema social particular.

Resumen

La cooperación es uno de los rasgos del comportamiento que definen a los seres humanos, así como a otros sistemas complejos, y que nos ha permitido evolucionar como especie. Sin embargo, incluso hoy en día, después de siglos de avances científicos, todavía estamos tratando de entender por qué cooperan algunos sistemas, y particularmente por qué lo hacen los humanos. Los experimentos de comportamiento humano, que se llevan a cabo en laboratorios y se basan en dilemas sociales modelados como juegos de comportamiento, tratan de arrojar luz sobre estos aspectos aún desconocidos.

Una de las aproximaciones más recientes ha sido el tránsito de la experimentación en comportamiento humano desde laboratorios hacia espacios públicos, donde los comportamientos ocurren de manera natural, para así estudiar los principales rasgos de comportamiento: cooperación, confianza, reciprocidad, aversión al riesgo o sentido colectivo. Las prácticas de ciencia ciudadana han proporcionado un marco perfecto para experimentar en el campo y así promover la participación de personas que generalmente no están involucradas en la práctica científica.

Esta disertación se centra en avanzar en el campo de la experimentación conductual en entornos abiertos utilizando experimentos enmarcados en las prácticas de ciencia ciudadana. El trabajo se divide en dos bloques, uno orientado al diseño de plataformas experimentales y otro centrado en el análisis de resultados experimentales. En el primero, concretamente, estudiamos cómo diseñar plataformas de ciencia ciudadana que permiten realizar actividades científicas en entornos abiertos y presentamos dos de ellas que se aplican en dos contextos diferentes. En el segundo bloque, realizamos estudios experimentales de sistemas sociales para analizar rasgos del comportamiento, buscar la emergencia de patrones de comportamiento y en última instancia evaluar los diseños de las plataformas.

Con la primera plataforma investigamos cómo los sistemas de ciencia ciudadana pueden servir como catalizador para la promoción del pensamiento científico y para aumentar la participación en actividades científicas. Presentamos Natural Patterns, una plataforma que presenta actividades científicas basadas en el método de investigación científica. Presentamos los resultados de un estudio de experiencia de usuario y, a la luz de las contribuciones de los participantes, describimos una serie de principios de diseño que aumentan la motivación de los participantes y promueven la participación en actividades científicas en el entorno natural.

La segunda plataforma está diseñada para estudiar rasgos del comportamiento humano y ayudar a crear experimentos en entornos abiertos, donde surgen fenómenos sociales reales, y lo hacemos fomentando la participación en el marco de la ciencia ciudadana. A tal efecto, la plataforma es modular e incluye una amplia gama de juegos de comportamiento. Es ligera, lo que permite la experimentación en lugares con poca infraestructura (espacios públicos, festivales o conferencias) y escalable, lo que permite introducir nuevos dilemas e interacciones sociales. Evaluamos la participación, la solidez y la calidad de los datos recopilados en todos los experimentos que se han llevado a cabo hasta el momento.

Con respecto a los estudios experimentales, presentamos los primeros dos experimentos de comportamiento humano que se llevaron a cabo con la plataforma. En el primero, los participantes deben predecir los movimientos del mercado financiero en diferentes circunstancias. Específicamente, estudiamos si las estrategias que emergen son sólidas en los experimentos réplica que se han realizado en diferentes lugares con muestras sociodemográficas muy diferentes. El segundo experimento estudia los patrones de comportamiento (o fenotipos) que surgen de las decisiones que toman los participantes cuando se enfrentan a un conjunto de dilemas sociales. Concretamente, nos centramos en el análisis de datos mediante técnicas de “unsupervised learning” y concluimos que el comportamiento de los participantes que toman decisiones sobre dilemas sociales con diferentes tensiones se puede caracterizar en un total de cinco fenotipos.

Los dos últimos experimentos presentados en esta disertación también abordan y profundizan en la experimentación del comportamiento humano. El primero es un “collective-risk dilemma” en el cual un grupo de participantes debe contribuir con sus activos para superar un daño colectivo, concretamente el cambio climático. Estudiamos cómo las desigualdades de recursos causan comportamientos injustos, provocando que los participantes más vulnerables se vean más perjudicados que los más favorecidos.

Finalmente, el último experimento es un “collective-risk dilemma”, pero en este caso no hay desigualdades económicas entre los participantes. A diferencia del resto de experimentos, este se lleva a cabo dentro de un colectivo determinado: el ecosistema de la salud mental, formado principalmente por personas afectadas por trastornos mentales, cuidadores y familiares. Estudiamos las tensiones que existen cuando todas ellas colaboran colectivamente para resolver un daño colectivo incierto. A partir de los resultados del experimento, se infiere que el coste de las acciones colectivas recae sobre los afectados por trastornos mentales.

En conclusión, las plataformas que hemos desarrollado nos ayudan a llevar a cabo actividades científicas participativas en el marco de la ciencia ciudadana, que tiene aplicaciones que van desde el aprendizaje hasta el activismo. Esto es especialmente notorio en el caso de la plataforma de experimentación que estudia el comportamiento humano in vivo, y que hace posible acceder a una muestra representativa de la población y a integrar los participantes en el proceso de investigación. La plataforma ayudará a establecer las bases para futuros experimentos de comportamiento en entornos abiertos. A la luz de los resultados de los experimentos sociales podemos comprender cómo nos comportamos colectivamente, en sociedad o dentro de un determinado grupo, cuando nos enfrentamos a dilemas sociales, y consecuentemente evaluamos los rasgos de comportamiento y la aparición de patrones de comportamiento gracias a técnicas de “unsupervised learning”. Estas ideas ofrecen nuevas formas de estudiar las tensiones que existen en una comunidad determinada en torno a un problema social particular.

Motivation

Social Experimentation and Citizen Empowerment

Cooperation is one of the three pillars of evolution, along with mutation and selection. Cells, multicellular organisms, animals, humans, organizations or corporations; they all cooperate. Cooperation is a defining feature of our human species; necessary, to some extend, to evolve (Nowak, 2006b; Axelrod and Hamilton, 1981a). Two classic cases that illustrate cooperation, particularly animal cooperation, are the cooperative transport of ants, which are capable of carrying very heavy loads, or the cooperative hunter behaviour of orcas.

In some cases, cooperation implies a benefit and a cost, a tension between individual and group interests. Among humans, those tensions emerge when, from the interaction and cooperation among individuals (or groups), one benefits while the other pays a price for that benefit to a greater or lesser extent. Social dilemmas reflect the tensions that arise from these interactions and game theory offers analytical formality (Van Lange et al., 2013; Kollock, 1998; Myerson, 1997; Dawes, 1980).

When the three Apollo 13 astronauts had to return to Earth, the cooperation between them was fundamental to the success of the spaceflight; the outcome was the same for the three of them, so there was no tension. Tension emerges when the outcome for the members of the group is not the same, a classic example being the lighthouse case. The lighthouse, and its light, is a public good that every boat that navigates the area benefits from. In general, lighthouses are built with public resources, therefore all citizens pay for them. Tensions emerge when we ask ourselves: Do we all make the same use of the lighthouse? Should we all pay the same? Either way, a community will be able to build a lighthouse though cooperation, and everyone will benefit from its light, independently of their contribution. In the extreme case, when tension is so strong that there is no margin for cooperation, there is only competition. We are then in front of a zero-sum game. Rock-paper-scissors game, chess or, according to some, politics, can be considered totally competitive games.

Cooperation (or its antithesis, competition) is just one of the main behavioral traits that can be studied. The study of human behavior sheds light on multiple behavioral traits (reciprocity, trust, fairness, coordination, etc.) that describe how we interact in society and therefore have a real impact on collective decision making. Nowadays, human behaviour researchers are immersed on specific questions with clear impact on social issues and, by extension, in policy making. Some examples of the study and implications of human behaviour could be found when people face decision-making tasks, collaborate to fight against a common bad (e.g. climate change) or to build a common good (e.g. social justice) (Tavoni et al., 2011; Milinski et al., 2008). The study of the sense of community, the relationships in an ecosystem or a particular social group (such as those affected by mental health or – war or climate – refugees), or ultimately studying the social capital of a particular human collective help to understand how social interactions are built (Eriksson, 2011; Pretty, 2003; Adger, 2003).

Social computer science studies social sciences, and particularly human behaviour, with techniques from other disciplines (e.g. maths, physics or computer science) that enable the analysis of large amounts of complex social data (Sánchez, 2018; Beuscart et al., 2016; Mann, 2016; Conte et al., 2012; Giles, 2012; Lazer et al., 2009). Science has evolved from an empirical science based on observation and description, to a theoretical science of patterns and generalizations, to today's computational science, which allows the simulation of complex phenomena. Nowadays, science explores and tries to obtain knowledge from large amounts of data generated and collected passively on multiple digital platforms, and because of the pervasive instrumentalization of human behavior, social science research has been transformed.

From the interactions between individuals emerge complex social dynamics. The study of human behavior can be seen as a complex system in which individuals are studied analogously to particles in physics, whose activity can be described by a few well-understood rules. Thus, studying how the system works allows us to understand how we behave in society. Obviously, individuals do not always behave according to those rules, but the rules would allow for a pretty good approximation to reality.

Experimentation has been necessary for the study of human behavior in order to improve theoretical models and simulations. It has traditionally been conducted in brick-and-mortar laboratories, with very particular characteristics in order to investigate specific aspects of the interaction between individuals or social phenomena in a scientifically rigorous manner. Currently, it is possible to perform these experiments outside the laboratory, much more efficiently and on different scales thanks to experimental platforms that provide versatility to run experiments in a distributed way (Radford et al., 2016).

Human behaviour experiments need diverse people for experimentation. This opens up an opportunity to involve the general public in scientific processes and serves as a gateway for them to participate not just as volunteers or experimental subjects, but as an active actor of the scientific research process. The idea behind this is encouraging citizens to participate actively in the study of human behavior, shedding light on how we relate (e.g. cooperate or compete) but also on the design of participatory experiences that improve citizen engagement in science.

Citizen science (Bonney et al., 2014; Gura, 2013; Hand, 2010; Silvertown, 2009) offers the perfect framework to introduce the general public in real research process and to stimulate the scientific action. Citizen empowerment based on knowledge, specifically on scientific knowledge, is essential to foster critical thinking and provide decision making tools to citizens. Opening science up to everyone outside laboratories helps to perceive it as part of a society's culture, reducing the gap between science and society. When an outsider of the scientific context must become a part of it, we, as researchers, have to start thinking not only about the scientific outcome, but also about the participant's experience in terms of learning or engagement (Preece, 2016; Bonney et al., 2015; Franzoni and Sauermann, 2014). Design based on the participant's experience and the study of participation in multiple fields (scientific disposition, learning, engagement, etc.) is completely necessary to engage with people.

There is a need to leave the "aseptic laboratories" and move the experiments to real contexts in which the participants actually interact. Therefore, another prerequisite has become imperative: the creation of platforms designed to collect experimental data with reliability but also ensuring that the experiments are attractive and capture the attention and interest of the participants, offering incentives that go beyond money, the classical catalyst for all behavioral experiments. In addition, all phases of the research process can be opened to the public, co-creating and co-designing with the scientific team. In this way, if the work has been created with the support of local communities, the impact of the research can be directed beyond science and towards objectives that have a social impact in the short-term.

In the same way that we, as social beings, are permanently connected and, therefore, receive influences from different sources, this work gravitates around several centers and emerges from issues that combine diverse disciplines which are connected.

One of the pivotal matters of this dissertation is the understanding of several aspects of human behavior. We study how we behave when facing social dilemmas, particularly when there are multiple levels of tension (harmony, stag-hunt, snowdrift, prisoner's dilemma). We dig into the study of collective behavior, how we deal with a global problem, specifically climate change, and the degrees of involvement, collaboration and risk that individuals take. We propose a set of games to study the behavioral traits within a specific context, in this case mental health, identifying the roles that different actors of the ecosystem have. In this line, an ambitious and highly motivating objective is to be able to create a more efficient interaction scenario at the social level based on the knowledge acquired about the behaviour of the system.

We also focus on the development of a platform that allows experimentation outside the laboratory. The need to experiment in human behavior in more real environments, where social relationships are more natural, means that the platforms need to have different characteristics than those traditionally built in laboratories. The challenge is to create friendly platforms for the participants, who are randomly recruited on the public space. The fact that we are in the field, in a technically hostile scenario, makes it necessary to build a tuneable and scalable system.

In this work, the experimentation in the field is performed following citizen science practices, specifically as pop-up experiments (Sagarra et al., 2016). The idea is to use citizen science to transform the participants from "experimental subjects" to "scientific experimenters". In this vein, the participants are part of the experiment by contributing with their actions to the behavioural games, but they participate beyond that, especially afterwards (due to the rigorous experimental protocols), by receiving feedback on the objectives and actions that are intended with the scientific activity, and in particular its social impact.

In this dissertation, in general, beyond the scope of behavioural experimentation, we try to understand how to build scientific experiments which promote the motivations of being part of science, help to understand the research method, to develop scientific thinking and advance in a change of attitude and disposition towards science. Considering all of this, we expect to introduce citizens in science and scientists in society. The call to scientific action must be driven by scientists, but it must go further and become a (scientific) social action, which puts everyone under scrutiny (primarily the scientist).

Plan and structure of the thesis

The thesis is organized following a collection of papers that have been partially or completely reproduced – depending on the contribution to each one – to avoid repetitions and gain consistency.

- **Part I: Introduction**

  Chapter 1 provides the foundations, a presentation of the main concepts in which the work is grounded: experimental game theory, citizen science and data analysis.

- **Part II: Platforms**

  In this part we describe two platforms of participation in scientific activities.

  Chapter 2 presents a study of HCI of an application designed and developed during a research stay at the Delta Lab at Northwestern University (Evanston, Illinois. USA.). We introduce an entirely new conceptual approach to designing citizen science systems, which serve as a catalyst to promote scientific reasoning and to increase the engagement in science (Vicens et al., 2018).

  Chapter 3 introduces a social experimentation platform with which we collect data of participants with the goal of studying human behavior. It is a modular platform that can conduct experiments based on social dilemmas. The chapter explains all the experiments carried out and the complete set of layers that the platform integrates to fulfil each experiment's requirements (Vicens et al., 2018).

- **Part III: Experiments**

  Chapter 4 presents Mr.Banks and Dr.Brain, the first experiments performed with the platform, which measure different aspects of human behavior. The first is focused on the study of decision-making and the second on the study of behavioral phenotypes when we face the dyadic games in the TS-plane (Poncela-Casasnovas et al., 2016).

  Chapter 5 introduces the experiment Dr.Brain, The Climate Game. This experiment is a public good games framed in climate change, also known as collective-risk dilemma. We study how participants collectively contribute to fight against a collective bad, in this case climate change, and how inequalities, forced into the game, provoke different participants behaviors (Vicens et al., 2017).

  Chapter 6 presents the collective risk dilemma within a specific group, namely those affected by mental disorders. In this case, we study how the members of the group (mainly of patients, families and caregivers) collaborate together for a common goal (Cigarini et al., 2018b).

- **Part IV: Concluding remarks**

  Chapter 7 summarizes the main contributions of the work and briefly discusses some future lines and open questions.

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Foundations

SUMMARY – Foundations introduces the basic concepts of game theory, citizen science and data analysis, which are the fundamentals of this dissertation. First, we present the social dilemmas that are the basis for the approximations to experimental game theory, as well as their application in social sciences. Secondly, we present citizen science, the approach taken to perform scientific activities, and particularly the experiments in human behaviour. We describe citizen science concepts, typologies, applications and impacts beyond the scientific outcomes. Finally, we focus on some methodologies of data analysis, mainly unsupervised learning, that have allowed us to study the behavioural patterns that emerge from the experiments collected.

1.1 Introduction

The foundations on which this work is based are constructed by border areas between disciplines. This makes it a multidisciplinary work in which different lines of knowledge converge. Therefore, to introduce the main ideas on which this thesis is based, it is necessary to look back and describe a set of basic concepts that arise from diverse disciplines such as mathematics, experimental economics, computer science, engineering, sociology or education, to give some examples.

First, we describe some basic concepts of game theory. Game theory studies the mathematical models of conflict and cooperation, and, for this purpose, it proposes a set of games with incentives. From this perspective, experiments have been carried out with subjects who have been confronted with a set of social dilemmas that conflict their interests. By studying their interactions, we gain a better understanding of how we make decisions and how we behave. In this first chapter we mainly introduce the mechanisms of the games with which we have experimented (a set of dyadic games and public goods games). The implementation of the experimental platform is described in Chapter 3 and the experimental results in Part III.

Public participation in science, whether in experiments or in other types of scientific activities, has become an area of study by a part of the scientific community under the name of citizen science. We introduce the concept of citizen science and what it represents. The participation of citizens in science forces us to rethink the way in which we contribute in science involving the general public. From our perspective, we focus on the design of technological platforms that allow the participation in a friendly way (e.g. gamification), emphasizing the intrinsic and extrinsic motivations for engaging the participants (e.g. learning, entertainment), and co-creating so that the participants feel part of the scientific process. We introduce concepts, typologies and projects behind citizen science that are reflected in the platforms presented in Part II and whose methodology is present in the experiments of Part III.

Finally in the last part we introduce the treatment of the data that has been collected in human behavior experiments. Mainly we will focus on the techniques of data analysis that have been incorporated into the study of social data, especially thanks to the computational social science, and that have been traditionally applied to other types of disciplines. In this case, we focus on some machine learning techniques, specifically unsupervised learning (i.e. clustering) and supervised learning (i.e. classification) applied in the analysis in the Part III.

1.2 Game theory

We humans are social beings, social animals. We relate to one another, and from the interaction between us emerge out dilemmas and conflicts, therefore behaviors (e.g. cooperation, collaboration, coordination...). But not only humans interact, non-humans also do so. Social agents (individuals, companies or governments), biological agents (bacteria, plants or animals) or technologies (algorithms, robots or networks) are adaptive agents, non-deterministic in their behavior. Game theory tries to understand systems formed by those agents studying the mathematical model to shed light into the adaptive agents' behavior.

> Game theory is "*the study of mathematical models of conflict and cooperation between intelligent rational decision-makers*" (Myerson, 1997).

Game theory handles social tensions in which two or more adaptive and interdependent agents make decisions based on their own motivations and the prediction of the other's decision. It applies to a wide range of scientific fields, especially economics, political science, sociology, biology or computer science. In this work we study decision-making and particularly social dilemmas; which is the study of the tension between the individual and the collective rationality (Kollock, 1998).

Von Neumann and Morgenstern provided the beginnings of the interdisciplinary research field of game theory in 1944 (Von Neumann and Morgenstern, 1944). From then to now the advances on game theory have been vast, developing cooperative and non-cooperative game theory, in the latter case introducing the concept of Nash equilibrium, which is extremely important in order to analyze individual strategies (in 1994 John Nash was awarded, together with Selten and Harsanyi, the Economics Nobel Laureates for their contributions to game theory). The study of dynamic models and the evolutionary game theory, initially applied to biology, have been important achievements on this discipline, and game theorist as Thomas Schelling and Robert Aumann have been honoured with prestigious awards. The implications in a wide variety of social systems, such as stock markets, ecosystems or political negotiations, have transformed game theory, and particularly behavioral studies, into a valuable area of study.

1.2.1 Nash equilibrium

John F. Nash defines in his work *Equilibrium points in N-person games* (Nash, 1950) the concept which came to be known as Nash equilibrium applied to non-cooperative games involving two or more players, based on the idea of equilibrium in physical systems. Nash equilibrium is an steady-state in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing his own strategy unilaterally.

Nowak (2006a) proposes to illustrate this concept with the following generic situation:

Consider the general payoff matrix:

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

There are two strategies A and B, so in order to find the Nash equilibrium we have the following criteria:

(i) A is a strict Nash equilibrium if a > c

(ii) A is a Nash equilibrium if a >= c

(iii) B is a strict Nash equilibrium if d > b

(iv) B is a Nash equilibrium if d >= b

Nash equilibrium is an equilibrium concept of non-cooperative games, whereas Pareto is a notion of optimality or efficiency originated in the economic equilibrium and welfare theories. Pareto optimality takes place when the resources in a system are distributed in the most efficient manner, it has no relationship with equality or fairness. In game theory we could say that it is a state in which no player can become better off without making someone else worse off.

An example to illustrate this situation could be the negotiations in climate change. All governments and citizens enjoy the benefits of a stable climate situation. Rich countries, such as China or United States, have been historically the most important greenhouse gas emitters due to the industrialization, thanks to which they progressed economically. Whereas poor countries produce low levels of emissions and are not industrialized, thus they do not progress. We are in a unfair situation, which makes it increasingly necessary to reach an international agreement with the largest economies, requiring them to cut-back emissions in order to protect the environment and, at the same time, do not harm developing countries further.

Simplifying the situation, two countries are negotiating to carry out massive cutbacks in greenhouse-gas emissions. In the first situation, both countries reduce their emissions, in that case they protect the environment and also reduce their productivity. However, on the basis of this situation, one country can be tempted to not reducing emissions and the other does it, consequently, the situation two is the one that reduces its emissions holds back on its developments but benefits both countries environmentally, the other continues producing emissions and developments. Nevertheless, both countries are tempted to not decreasing emissions. Finally, situation three, both countries do not decrease emissions; as a result both countries are affected environmentally but they maintain their productivity.

Nash equilibrium here represents the third situation. If one country changes the strategy it would be in the second situation, it would be reducing while the other would not reduce. In this case, the Pareto optimality is the opposite, first situation: both countries reduce emissions and their productivity has a better effect, but also the second situation is Pareto optimal. The simplification of this case is the well-known prisoner's dilemma, explained below with more detail.

The payoff matrix of the prisoner's dilemma that illustrates the climate negotiation analogy could be:

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \end{array}$$

Following with the analogy, in the first situation the choices are (A, A), both win (1). In the third situation, the choices are (B, B), both lose (0). In second situation, the choices are (A, B) and (B, A), the one that chooses B win much (2) and the other lose much (-1).

1.2.2 Dyadic games

"Game" in game theory is any context in which agents (e.g. individuals) interact and the outcome of this interaction is related to the choices of all the agents. The payoff of one agent is associated with the interactions of the others; in other words, the choices and actions taken by one affect the others. In everyday life this is a very common situation and those games are analogies of (social) dilemmas that arise when individual and collective interests are in conflict.

We define a general form to represent the most common dyadic games and the most simple social dilemmas involving two people facing a single choice between two options, two-player two-strategy (Kollock, 1998; Rapoport and Guyer, 1966). Those games are the harmony game (HG), the snowdrift game (SG), the stag-hunt (SH) and the prisoner's dilemma (PD). The payoff matrix that represents those games is:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array}$$

Each player makes a choice between "cooperate" (C) or "defect" (D), and depending on their choices they receive a payoff. If both cooperate, the two players receive a reward (R); on the contrary, a bilateral defection implies a punishment (P); temptation (T) happens when you defect and the other cooperates; and finally, the sucker (S) occurs when you cooperate and the opponent defects.

We build the four dyadic games fixing R and P, and giving different values to S and T. The TS-plane is divided in four quadrants and each of them represents a game depending of the relative order of the payoff, such as: harmony game ($S > P$,$R > T$), snowdrift ($T > R > S > P$), stag-hunt ($R > T > P > S$) and prisoner's dilemma ($T > R > P > S$).

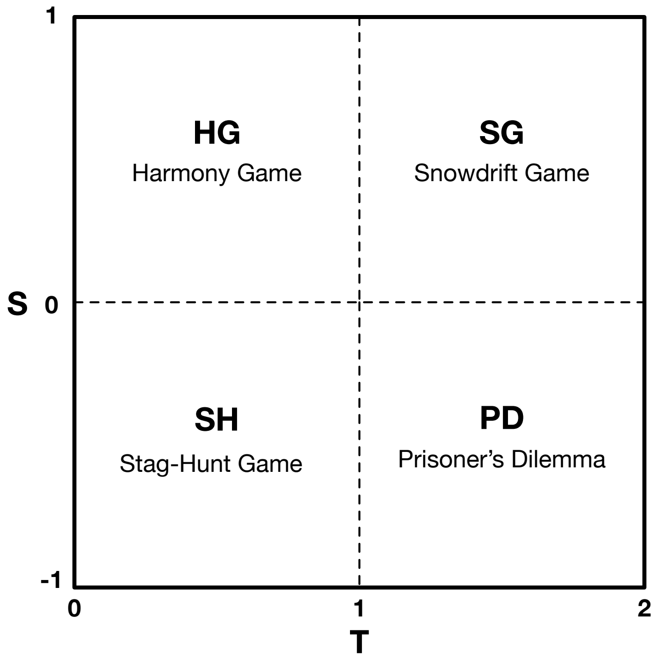

Figure 1.1: TS-plane dyadic games. Set of dyadic games in the TS-Plane. Prisoner's Dilemma, Stag-Hunt Game, Snowdrift Game and Harmony Game.

Harmony game

Harmony game ($S > P$, $R > T$) (Licht, 1999) is a trivial game that completes the TS-plane, taking values of $0 < S, T < 1$. The strategy here is unequivocal, the cooperation choice gets higher payoff regardless the choices of the other players. Cooperation strictly dominates defection, therefore the unique Nash equilibrium is that both players cooperate.

Snowdrift game

The game snowdrift, hawk-dove or chicken game (Sugden, 2004), is a model of conflict or anti-coordination between two players. The principle of the game can be drawn from the story that emanates from the name itself, chicken game, well-known thanks to the movie "Rebel Without a Cause". Two drivers race towards a cliff with the object of not being the first driver to slam on his breaks and, at the same time, not plugging over the cliff. The original story of the game is about two drivers that drive towards each other on a collision course in a narrow road: one must swerve, or both may die in the crash, but if one driver swerves and the other does not, the one who swerved will be called a "chicken", meaning a coward.

In this game, what stands out is the high temptation to defect and the low punishment, hence bilateral defection leads to the lowest payoff. When both players cooperate, both receive a reward; however, they have a strong incentive to defect, and in that case the payoff increases, so they have the temptation to defect, which is better for the one who defects and worse for the other. Otherwise, in the scenario where both players defect they receive a punishment, so it is more beneficial for both that at least one of them chooses cooperation. Finally the solution, Nash equilibrium, is that one chooses defection and the other cooperation (D,C) and vice versa (C,D), the two pure contingent strategy profiles. This is a anti-coordination game and there are not dominating strategies. It has three Nash equilibria, the third is a mixed equilibria in which each player choose with a certain probability between the two pure strategies.

Stag-hunt game

Coordination is a fundamental behavior in our society. Norms and conventions have evolved involving many factors (Young, 2015) and nowadays we are coordinated in many ways, following conventions like driving on the same side in a given country or speaking the same language in order to understand each other; or social norms such as not smoking indoors or flushing when you are done in the bathroom . Social cooperation is common in animals, which require it to achieve their goals; we can see it in the coordination of slime molds or the hunting practices of killer whales.

Stag-hunt ($R > T > P > S$) (JABONERO, 1953) is a coordination game originally explained as the dilemma between the cooperation with a group of hunters to kill a large stag, or deciding to individually kill a hare. A hare is worth less than a stag, however social cooperation is necessary in order to hunt a bigger prey.

The stag-hunt game has three Nash equilibrium, two pure and one mixed. Mutual cooperation and mutual defection are both pure. When both players cooperate, they achieve the bigger payoff, while if both players defect the strategy is less efficient but also less risky. The third is a mixed strategy with limitations in their equilibrium.

Prisoner's dilemma

The prisoner's dilemma (Rapoport and Chammah, 1965; Axelrod and Hamilton, 1981b) is the paradigm of the study of cooperation. It was proposed by Merrill Flood and Melvin Dresher in 1950 and reformulated with the penitentiary theme for which it is well-known by Albert W. Tucker. It is presented as follows:

Two members of a criminal gang, Maria and Toni, commit a crime and are arrested. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors lacks sufficient evidence to convict the pair on the principal charge. The authorities plan to sentence both to a year in prison on a lesser charge. Simultaneously, each prisoner is given the opportunity either to betray the other, by testifying that the other committed the crime, or to cooperate with the other, by remaining silent.

The offer is:

- If Maria and Toni each betray the other, each of them serves 2 years in prison.
- If Maria betrays Toni but Toni remains silent, Maria will be set free and Toni will serve 3 years in prison (and vice versa).
- If Maria and Toni both remain silent, they will only serve 1 year in prison (on the lesser charge)

The prisoners will have the opportunity to reward or punish their partner and their decision will not affect their reputation in the future. Betraying a partner offers a greater reward than cooperating with them, and purely self-interested prisoners would betray the other, so the outcome for two purely rational prisoners is for them to betray each other. However pursuing an individual reward logically leads both prisoners to betray the other when they would get a better reward if they both cooperated (Kennard, 2015).

The dominant strategy in prisoner's dilemma ($T > R > P > S$) is defection. It always results a better payoff, and mutual defection is the only Nash equilibrium. However, this choice does not correspond to a socially optimal outcome. Pareto optimal describes the social optimal, in this game all the options are Pareto optimal except for the Nash equilibrium, which is defect in both cases. All of the overall efficient outcomes are the ones that do not occur in equilibrium.

Dictator game

Dictator game (Guala and Mittone, 2010) is probably the simplest experimental setting one can think of. Sometimes it is not presented as a game because the interaction is minimal. The idea of the dictator game is very simple: two players are to divide a sum of money provided by the experimenter. Only one player, the dictator, can determinate the proportion of the shares, and the other can only accept the proposed division. The Nash equilibrium of this game, based on the assumptions of rationality and self-interest, is for the dictator to keep all the money for himself, however the experiments in lab reveal that the behavior is not rational.

Another version of this game introduces a new player whose objective is to punish the decision of the dictator with a certain cost. Same as in the standard version of the game, the dictator splits money between him and the other player. In this case, however, an observer, the punisher, supervises the division and has the possibility to punish the dictator using his own money. The Nash equilibrium is for the dictator to keep all the money, as well as for the punisher to not punish and also keep all of his money.

Trust game

The trust game (also known as the investment game), which was designed by Berg et al. (1995), is an experiment that measures the trust in economic transactions. Berg et al. propose the game as follows:

In stage one, two subjects, Maria (with the role of investor) and Toni (an anonymous entrepreneur), are separated in two rooms and are each given a quantity of money as a show-up fee. Maria must decide how much of her money to send Toni . The amount sent is then multiplied by a factor (commonly tripled). In stage two, Toni the entrepreneur must decide how much money to return, which is denoted.

The Nash equilibrium in trust game requires the entrepreneurs to not give any money back to the investor, so the only Nash equilibrium sees the investor keep the money for himself. However, in several experiments (Berg et al., 1995; Brülhart and Usunier, 2012) the actual results in both ways, trust and reciprocity, differs sharply from the predicted results under the assumption of pure self-interest.

1.2.3 Public goods game

Group cohesion and social norms are some of the sociological interpretation of the public goods game. It puts the personal and the public goods in tension, in a similar way than the prisoner's dilemma. In fact, public goods game is a generalization of prisoner's dilemma to an arbitrary number of players (Dawes, 1980).

Typically the experiment is proposed as follows:

An experimenter gives an amount of money (e.g. 20 MU's) to each participants (e.g. 6 players). The participants may contribute part (or all) of their endowments to some common pool. Then, the experimenter multiplies (by a factor greater than 1 and lower than the number of players, e.g. 3) the amount accumulated in the pool and divides it equally among the players, irrespectively of the amount contributed by each individual.

In the case of all the players contributing maximally, they end up with the maxim possible outcome equally shared. However, each individual is faced with the choice to behave as a free-rider while the others contribute, which means that a high temptation to exploit exists. Therefore, the dominant strategy or Nash equilibrium is to invest nothing at all (Hauert et al., 2002). In actual experiments players tend to invest a lot (Fehr and Gächter, 2002).

Public goods game represents numerous situations in real context (climate negotiations, investment in public policies, contributions to private parties...), in which the cooperators are tempted to act as free-riders, along the same lines of the prisoner's dilemma but introducing more actors. In fact, in some experimental settings the game is not played neutrally but is framed in real-life situations. Public goods also have different variants introducing inequalities in the players' endowments, iterated decisions, punishment and rewards or minimum threshold to return the inversions, among others.

Experimental game theory

Experimentation in game theory has been a relatively new approach to the study of human behavior and decision-making. In fact, experimental economics were almost non-existent until the mid-1960's (Levitt and List, 2007b). Crawford (2002) refers to experimental game theory as experiments whose goal is to learn about general principles of strategic behavior and notes that experimental game theory has been questioning and modifying the theoretical game theory. The empirical orientation of game theory has turned experiments into an important tool for the analysis of strategic behavior. Crawford, V. P. also remarks the need for empirical information about principles of strategic behavior and the advantages of experiments in providing it as two main factor in the emergence of experimental game theory.

The experimentation adds context effects that is not possible to introduce in theoretical approaches. There are several exemplifications of that, for instance the effect of social distances in the experimental dictator's game (Bohnet and Frey, 1999) or windfall versus earned money (Cherry et al., 2002) or factors that in consequentialist theories are irrelevant as shows Falk et al. (2003) in ultimatum games. Therefore the behavior in experimental games is sensitive to factors that, in principle, cannot be explained, or even studied, by theory.

In some cases, controlled experiments results have contradicted theories well supported theoretically, as is the case of the homo economicus model. Richard Thaler, who won the Nobel in Economics for "killing" the homo economicus, demonstrates that mankind make decision afflicted by emotions and irrationality (Thaler, 2000; Gintis, 2000) in the same way that Robert Shiller won the Nobel in 2012 showing that the markets are not rational. Markets bounce up and down more than could be justified on the basis of rational fundamentals as it is concluded in the book *Market Volatility* (Shiller, 1992) that represents a culmination of his research in this topic.

The experimental method to study physical science applied to the understanding of human behavior has been questioned. There is a traditional concern about experimentation in labs and about the "realism" and the principle of generalizability, the insights gained in the lab extrapolated to the world beyond (Levitt and List, 2007b). However there are a lot of indicators that emphasize the benefits to realize experiments and confront it to the theory (Cesarini et al., 2009; Falk and Heckman, 2009).

Experimentation has an enormous importance in the theoretical fundamentations. Laboratory experimentation allows for the use of the data to conduct simulations and improve models of behavior. The coexistence between theory, experimentation and simulation is fundamental to advance in understanding of human decision-making.

Nowadays, behavioral experimentation has other alternatives to perform "lab" experiments that allow collecting data from multiple samples, carrying out multiple interaction and large populations, precisely moving the-labs-to-the-field (Sagarra et al., 2016), living labs or creating online-labs (Holt, 2018; Fischbacher, 2007; Radford et al., 2016; Chen et al., 2016). Our work is developed under this premise, performing behavioral experiments to-the-field and building the tools and platforms to carry them out.

1.3 Citizen Science

Science is moving forward. The number of publications is doubling every nine years (Bornmann and Mutz, 2015) but a proliferation of papers does not always imply a growing of knowledge and that knowledge is not always transmitted to general public. Organizations, policymakers and citizens are disconnected from scientific activities and knowledge, and there is at least an opinion gap between scientist and the general public, (Leshner, 2015) who debates about ideas on which scientist have already agreed. This disconnection makes a public understanding of science difficult .

The scientific community is responsible for the creation of a better communication and for the promotion of a greater involvement of the general public in science. Science has an impact on society, at least to help taking better decisions. Engaging people in science, opening a public debate beyond the scientific world, clarifying the real impact of research, the utility of science in society and the motivations of scientists are all fundamental to open science to society (Nature Editorial, 2017b).

A collaboration between scientist and the general public has the potential to expand the understanding of science, as well as contribute actively in research with new ways of looking at and thinking about science. This new approach has positive consequences and outcomes if the initiative makes convenient protocols available to citizens , in addition to controlling and measuring the quality of data (Bonney et al., 2014).

Citizen science tries to open science to a broader public, offering the possibility to participate actively in a real scientific research. Citizen science has been defined in multiple ways.

> Oxford English Dictionary defines citizen science as "*the scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions*" (Oxford University Press, 2014).
>
> SciStarter, an online platform to participate in scientific projects through formal and informal research, defines citizen science as "*the public involvement in inquiry and discovery of new scientific knowledge. A citizen science project can involve one person or millions of people collaborating towards a common goal. Typically, public involvement is in data collection, analysis, or reporting*" (SciStarter, 2017).
>
> Socientize defines citizen science as "*the general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources*" (Consortium, 2013).
>
> Or Cohn, talking about the possibility to do real research with citizens, says "*citizen science is a form of research collaboration involving members of the public in scientific research projects to address real-world problems*" (Cohn, 2008).

Science, historically, has been performed by non-professionals, people who was not paid as scientist. The production of science became a profession in the late 19th. Before that, the majority of scientific contributions were made by amateurs, amateurs but experts (Miller-Rushing et al., 2012).

During the scientific revolution (17th and 18th century), a time of splendor in science, when the scientific academies were founded and developments in mathematics, astronomy, natural sciences and other disciplines transformed society and defined the basis of modern science as we know it today, the majority of scientists were amateurs, such as Benjamin Franklin, Darwin (Silvertown, 2009) or even Einstein, who was working in a patents office when he published the special theory of relativity in the early 20th century (Einstein, 1905).

AsSilvertown explains, amateur scientists' contributions were traditionally seen in disciplines like astronomy or natural science, in which observation is more important than equipments. Nowadays, modern citizen science is a bit different. A professional scientist designs a context that promotes participation among the public, not only providing tools but also designing activities, studies or experiments with the aim of granting an active role of the participants in research.

Diverse terminologies are used to refer to projects in which volunteers contribute in science research, such as crowd science, civic techno-science, street science, and so on. Cooper et al. (2014) suggest, given the invisible prevalence of citizen science, to use consistent terminology in order to become acquainted with the impact of citizen science research . A study conducted by Kullenberg and Kasperowski (2016), in which they analyze two datasets from Web of Science using a wide variety of descriptors, illustrates that in the last ten years a number of publications related to citizen science have increased from 41 publications in 2005 to 402 publications in 2015. The same study sheds light on the broad spectrum of disciplines in which citizen science can have applications, such as ecology, environmental sciences, computer science, or geography, and. this transdisciplinary approach interferes in the process of measuring the impact of citizen science.

Technology brings science closer to citizens, it stimulates follow-up participation and helps to play and active role in science. Citizen scientist have opportunities to join scientific projects launched by scientist in different fields and with different roles. The proliferation of personal devices with GPS, sensors or cameras allow citizens to collect, contextualize and share data with scientific purposes in the wild. Detractors are concerned about the data in aspects like quality, privacy or biasing. Stimulating citizens to take part actively in science activities is positive, but it is also important to shed light on the participants' motivations concerning conflicts of interest (Editorial, 2015).

The proliferation of research using the citizen science approach has generated scientific local and international conferences and the creation of communities to support citizen science activities with the aim of collaborating and sharing knowledge. In 2012, the Citizen Science Association (CSA) took shape in the USA, and the European Citizen Science Association (ECSA) was launched in 2013 due to a growth of the Citizen Science movement in Europe. Also, theAustralian Citizen Science Association (ACSA) was founded in 2014.

The federal government of the United States of America promotes open innovation, and citizen science is a particular case of it. In 2009 (The White House, 2009) and 2013 (The White House, 2013), the Obama administration encouraged agencies that took actions to support transparency, public participation and collaboration by sharing data sets. The number of federally-sponsored citizen science projects increased thanks to these measures, with positive impacts not only in scientific terms but also in public policy (Bowser and Shanley, 2013).

Definitively, in recent times open minded scientist, citizens and policymakers have been working together to create open science for the broad public with an impact in society thought active participation, collaboration, the sharing of knowledge and transparence.

Typologies, projects and outcomes

There is a broad range of ways to classify citizen science projects based on different characteristics. Ceccaroni et al. (2017) review the most recent classification of the projects. This classification is not mutually exclusive; the same project can belong in different typologies, because it is classified under different parameters. Cooper et al. (2007) classify the initiatives by the different ways of collaboration between professional and non-professional scientists; Danielsen et al. (2009) define the projects based on the specific public participation in certain domains; Wiggins and Crowston (2011) are focused on the degree of public participation; Shirk et al. (2012) classify the projects by the direction model and the public implication in the scientific process; Haklay (2013) also use the governance model and the participation and collaboration between professional and nonprofessional scientists; and finally Bonney et al. (2015) classify project by the nature of the activities participants engage in.

Socientize classifies citizen science in categories such as collaborative science, crowd-crafting, participatory experiments, collective intelligence, volunteer thinking, volunteer sensing, volunteer computing and human sensing (Consortium, 2013). The spatial levels in which citizen science works are from local to global in a real and virtual world.

The projects presented in this work belong to different categories, but all of them are data-collection projects, in which the participants contribute with data that can be used in research. The collection of data is performed in a contributory or collaborative way. In the contributory projects, the participants only collect data; however, in collaborative projects, they also help on other steps of the research projects such as the analysis of data, discussion results, dissemination of findings, etc. The setup of performing projects differs from crowdsourcing projects to participatory experiments. In the former, the contributors provision data remotely through online platforms, while in the latter the participants take part in an ephemeral experience in-person.

The first few projects in citizen science were about observation and collection of natural data. The Christmas Bird Count was launched in 1900 to collect data on birds (National Audubon

Society, 2017). The American Breeding Bird Survey (BBS), a long-term, large-scale, international avian monitoring program, was initiated in 1966 to track the status and trends of North American bird populations (USGS, 2017).

Traditionally, most of the citizen science projects study environmental sciences in which the nature observation and data collection is essential. Participation helps to increase the data collected and even makes it possible to reach places not available to researchers without public collaboration. All of these becomes much simple with the technology available today.

We identify a broad range of disciplines that habitually do not explore public participation but, paradoxically, those disciplines need data generated by people. Generally, the data is produced in very controlled scenarios, like labs, that allow a better reproducibility, or even in social networks or mobile devices, often from people who had not been directly informed. Nowadays, however, those disciplines have experienced an exponential growth led by researchers who believe open innovation and a direct connection with the public is highly positive and enriching in both scientific and social terms.

Citizen Science brings new opportunities to improve the interaction between science, policy and society. It is an important source of knowledge in those three levels. From the perspective of policy, citizen science projects could be designed from very different angles. A collaboration between citizens and organization at the local or national level shed light on how citizens face real problems, especially at a local scale, but also organizations can support projects with the aim of raising social awareness. The domains in which citizen science initiatives merge the needs of policymakers and citizens are diverse in their implications: urban planning, social awareness, smart cities, food or services. Citizen participation is performed with different typologies of activity depending on the aim of the contribution. Participatory science, environmental observation or passive sensing are some of the possibilities (Haklay, 2015). Generating a high interaction between citizens-scientist-policy makers is much more feasible if the projects deal with issues that affect them all. Citizen science serves multiple targets and it has implications for the three main agents. From a scientist perspective, it is crucial to contribute with scientific literacy and, as the activities are addressed to general public, to try to make the research compatible with social concerns. Contributions are essential in science; however, the educational purpose, social impact or science awareness are other required conditions of the uttermost importance.

The collaboration between all the agents has the potential to broaden the scope of research thanks to a fresh view from outside the academia, and also to improve the ability to collect significant data for scientific projects. The community members have better access to local data and they are learning about the central theme of the research (Cohn, 2008). Therefore, the potential benefits of citizen science projects go beyond scientific knowledge and literacy, it has an impact in policy, education, the participants and society at large (Raddick et al., 2010).

1.3.1 Participatory experimentation

Involving the public in scientific activities is fundamental in a lot of disciplines of research for different reasons. For instance, in social sciences the collaboration of volunteers has been necessary to perform experiments in human behaviour, cooperation, economics or social interaction. In fields like urban planning, environmental policy or even epidemiology, public participation endorse projects providing data which is usually hard to collect without the public involvement, and producing research results that could not otherwise have been achieved.

To carry out this research, it is necessary to collect data from different sources. For instance, in social science experimentation the participation of volunteers is necessary Traditionally, students have been used as subjects of studies without their total implication beyond contributing with data. In environmental science, for instance, getting samples of nature that are scattered geographically or in inaccessible locations require the participation of volunteers to have access to remote areas or get a composite sample.

Nowadays the growth of data available in online platform and devices, such as social networks, mobile phones or smart city applications, enables the access to large amounts of content, which is known as big data. Big data is described by the 3V’ĂŹs: Volume, amount of data; Velocity, baud rate; and Variety, heterogeneous data; and sometimes with a fourth V, Veracity, trust in data and outcomes. Big data does not include citizens in the research process, it just uses the data generated by users without a direct communication. The lack of knowledge, transparency and context surrounding the data source are a cause of concern and deep discussion not only due to privacy issues, but also due to the veracity of data and outcomes.

To produce research that includes an active involvement of the public, the orthodox mechanisms from science experimentation are not functional because the goal is not only a scientific outcome, but also a public return, hence both have to respond to social needs and scientific questions. The approach needs to foster participation, engage participants in the research and provide tools that allow the participation in different formats, such as in the wild or online, so the experiment design is based in principles of open participatory science.

Sagarra et al. (2016) define a particular case of participatory experiments, called: ĂIJA Pop-Up Experiment. It is a physically light, very flexible, highly adaptable, reproducible, transportable, tuneable, collective, participatory and public experimental setup for urban contexts that (1) applies Citizen Science practices and ideals to provide groundbreaking knowledge and (2) transforms the experiment into a valuable, socially responsible, consented and transparent experience to non-expert volunteered participants with the possibility to build the urban commons arisen from facts-based effective knowledge valid for both cities and citizens. The implementation of those experiments are halfway between classic behavioural experiments and big data.

1.3.2 Crowdsourcing

Formulated by J. Howe and M. Robinson in 2006, crowdsourcing refers to a new business model that resolves a business problem taking advantage of a network of collaborators through an open call for contributions.

> Howe defines crowdsourcing as "*[..] the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the wide network of potential laborers*" (Howe, 2006).

D. C. Brabham, author of the first academic article using the term crowdsourcing, links crowdsourcing with crowd wisdom; a collective intelligence that emerges from a group of people working in a problem’s solution as a result of the aggregated work (Surowiecki, 2004). All of this is motivated by technologies that facilitates the collaborative work (Brabham, 2008).

Crowdsourcing is used in broadly diversified applications in science, research, journalism, public policy, funding and so on. Nowadays, different categories of crowdsourcing are standardised in our society in projects like Kickstarter, a crowdfunding platform that collects funds for projects collaboratively; Scistarter, a platform that helps to contribute in citizen science projects; reCAPTCHA, an implicit crowdsourcing project through which archives of The New York Times or books from Google Books are digitized; Wikipedia, the popular online encyclopedia built with contributions of the crowd; or even Uber, that allows people to drive people elsewhere leveraging the algorithms that detect the nearest driver from the consumer’s location.

This way of working, which splits tasks between a group of people with the aim of completing a major activity collectively, was already adopted before the influence of digital technology in our daily life. However, the rise of crowdsourcing initiatives correlate with the rise technology in our daily life..

Focusing on science, all scientific activity is based on collaboration. The exchange of letters between natural philosophers in the seventeenth century develops the concept of Invisible college, a precursor of The Royal Society of London (Price and Beaver, 1966). From that time onwards the way in which scientists collaborate has been studied under different perspective, from sociology to complex networks. Now, more than ever, the collaboration in multidisciplinary research is essential, and the lone author has disappeared. From the seventeenth century to the early twentieth, one author per paper was the most common (Greene, 2007). However, the co-authorship model is now widespread and has increased over the years. Some disciplines like physics have been published papers with more than 3,000 authors (Adams, 2012).

Traditionally, the contributors in scientific projects are scientist and researchers, who are science professionals. Particularly in science, crowdsourcing introduces the general public to the the chain of scientific collaboration. In 2010, Nature published a paper with an unusual author: Foldit Players (Cooper et al., 2010). Foldit[1] is an online game in which the crowd tries to solve biological problems, specifically protein folding. Foldit is based on Rosseta@Home, a project of distributed computing that uses spare processor time to solve scientific problems. Foldit is designed as a game where users start solving simple structural problems that become increasingly difficult, with the objective to fold the structures of proteins. After that, the solutions with a higher score are studied by researchers. This is an example of a hybrid between crowdsourcing and distributed computation in which the gamers, non-professional scientist, are considered authors of an academic paper.

Another crowdsourcing project with high impact in the scientific literacy, with more than 50 publications, is Galaxy Zoo[2]. These publications cover a wide range of disciplines, such as space, physics, climate, humanities or medicine. Galaxy Zoo is a platform that asks the user to classify galaxies according to their shapes. For most of the twentieth century, classify galaxies based on their morphology has been a standard practice. The huge amount of galaxies to classify collected from surveys as Sloan Digital Sky Surveys (nearly one million) makes it difficult to visually inspect them by small teams of astronomers. Galaxy Zoo was launched with the purpose of dividing the work between multiple observers and increasing the confidence in the classification, in order to to reach a sufficient critical mass to analyze millions of images. Based on the idea of SETI@HOME and similars projects of volunteer computing, in this case the idea is to use volunteer's skills to classify galaxies. Approximately 100.000 participants made more than 40.000.000 classifications (Lintott et al., 2008).

Sauermann and Franzoni (2015) synthesised divers types of benefits of involving the crowd in scientific research. One of them is the motivation to take part in crowdsourcing, which is an intrinsic motivation rather than an economical one, meaning that the cost is low in human resources. Also, working in parallel tasks helps achieve the objective in a shorter time. The diversity of participants integrates skills not typical in research, and the project benefits from diverse profiles and background. Moreover, the crowd increases the space and time variables, covering more places during more time, and involving the general public has outcomes in education. However, their findings prove that the effectiveness of the projects is variable depending on the kind of project, and that most contributions are done by a small group of participants.

Two different approximations to crowdsourcing are developed in this thesis. First, a crowdsourcing initiative to study the participation in cultural events. The idea behind it is to collect

1. Foldit. http://fold.it/
2. Galaxy Zoo. https://www.galaxyzoo.org/

opinions about the experience of visiting museums as well as to study mobility patterns which provide information about the different spots visited by participants. Secondly, a project to shed light on the motivations and science disposition in the participation in science activities in the wild. In both cases, we collect information to better understand collective behaviour based on the individual one. The aim is to study the outcomes of participation in educational terms, science disposition and creation of collective knowledge as a reason to participate in these activities.

1.3.3 Motivation

Why do people participate in citizen science projects? Improving the strategies to recruit participants for scientific activities is an important goal in citizen science. A wide array of techniques are used to draw the attention of potential contributors, but motivation is a main requirement for engagement and it encourages ordinary people contribute actively in science. Traditionally, most of the research in citizen science was focused on the benefits for scientist and on the scientific goal. However, recently studies on citizen science participation shed light on why people participate in research and who is this people.

The different typologies of citizen science projects open a large variety of potential motivations depending on the task designed in each one. Zoo Universe, the project that classifies astronomical images, concluded, after studying the main motivations of their participants, that the most important motivations are: the topic (astronomy), the contribution, vastness, beauty or fun (Raddick et al., 1988). Malone argues that the motivations are money, love and glory (Malone et al., 2010). However, in citizen science projects the contributor does not usually receive any economic compensation. Rather than that, citizen scientist indicated that some of the motivations are an inherent interest in the subject of scientific inquiry, the perception that a activity will be fun and engaging, altruistic reasons or the interest in contributing in real research with experts (Prestopnik and Crowston, 2012).

Either way, there are two open fronts: recruiting and retaining ordinary citizens in citizen science projects. In this context, knowing the intrinsic and extrinsic motivations (Bénabou and Tirole, 2003), as well as knowing the research base around recruiting and retaining volunteers, it is crucial to achieve these objectives (West and Pateman, 2016).

1.3.4 Games

Everyone likes to play games. Concepts as gamification, playful design or serious games facilitate the grouping of techniques in the framework of ludification. The difference among them is based on the relationship between two axis: play/game and whole/parts.

> Gamification, was defined by Deterding et al. as "*the use of game design elements in non-game contexts*" (Deterding et al., 2011).

Gamification is a term which was coined in 2002 by Nick Pelling, and it has gained popularity years after. However, the idea is not new at all, since the 80's, game design research has been studying many aspects related with games, such as gamers behaviour (Greitemeyer and Osswald, 2010; Gentile et al., 2009), motivations (Deterding, 2012) or experience (Chen, 2007).

The influence of video games in our behaviour has been studied traditionally, sometimes from a negative perspective and focusing on violence. But, for us, the most important thing are the potential benefits, the positive perspectives to create prosocial games. Motivation is a crucial aspect to use games in non-context scenarios. We need to use game design to create games that motivate people to play, design focusing on the motivations. Finally, the user experience, game experience and playability among others have been also vastly studied in order to improve the design of the games and the experience of playing them. All the accumulated knowledge is applied to incorporate elements of games for the purpose of improving the engagement, becoming fully immersed (flow) or leveraging participation in activities that are normally not funny. The key idea is to use games as a means to educate, engage in science participation and address social issues such as equity, public health or environmental sustainability (Horn, 2014) to name just a few.

Creating game experiences in an ordinary context using game design principles is a good way to improve the quality of the whole experience. Because of that, a playful design is implemented in multiple applications such as: education (Prensky, 2003; Villafuerte et al., 2012), music (Denis and Jouvelot, 2005; Villafuerte et al., 2012), health (Wattanasoontorn et al., 2013; Villafuerte et al., 2012), energy saving (Banerjee et al., 2016) and also in science (Morris et al., 2013).

In citizen science, and particularly crowdsourcing, it is essential to motivate the crowd to take part in the activities. In order to maintain the motivation of participants, gamification techniques are very useful. Gamification, in crowdsourcing, basically increases the intrinsic motivation of the users, the elements of a game incentivize participation and have psychological (motivation, enjoyment, attitude) and behavioural outcomes (Morschheuser et al., 2016a).

The implication of volunteers in citizen science is a key point to measure the success of the project. Not only the scientific initiative needs to have scientific value; the volunteers should also be in tune with the value derived from the research. Gamification has the potential to maintain engaged volunteers and bring new participants (Bowser et al., 2013). However, as Bowser et al. explain, gamification has to ensure that it is not a simple marketing tool and that the participants really get benefits. Because of that, when designing and application it becomes to take into account the potential benefits for the users, motivations and engagement. The well documented benefits of citizen science are diverse, some of them are: educational, the change of attitudes and behaviour positively, or the gain of knowledge about community structure (CAISE Inquiry Group, 2009). Henceforth, applying gamification techniques in activities with a positive impact is a responsible and proper manner to incentive the general public participation.

Morschheuser et al. (2016a) defense the usefulness of gamification and its positive impact in citizen science initiatives. They remark on the positive effects of gamification, such as engagement and motivation or data quality, which are essential (Crowston and Prestopnik, 2013; Sheppard and Terveen, 2011). In terms of engagement, the individual motivation elements work well in short-term engagement while the social elements perform the best in long-term engagement (Lee et al., 2013); in this case four categories of intrinsic and extrinsic motivations have been identified: contributing in science, learning and personal interest, communication and entertainment (Tinati et al., 2017). The quality of data between gamified projects and non-gamified projects is not significantly different (Prestopnik et al., 2014; Cechanowicz et al., 2013).

Citizen science projects developed in this thesis, both crowdsourcing and participatory experiments, combine the citizen science basis with gamification techniques. In the crowdsourcing projects, developed as a mobile application, the idea is to engage people to participate and maintain a long-term involvement. However, in the participatory experiment the idea is to create an experience instead of an scientific experiment, but maintaining scientific rigorosity in the process.

1.3.5 Learning

School-based learning is only a (small) part of the educational needs of today's society. He experiences of everyday life in a wide range of contexts and situations are a great source of learning: homes, museums, communities, television, labs, research centers or in the wild. One of the most important citizen science objectives is to install scientific thinking into the everyday life.

Citizen science projects involve the public in scientific concepts that are relevant in our lives. The public understanding of science helps to comprehend the process of research, how the scientific community discusses subjects that affect our life directly or indirectly. Science deals with all the stuff around us, like technological advances, health, environment, human behaviour or natural hazards, among others. The participation in these projects provide a framework in which participants are able to understand and engage in scientific activities (Trumbull et al., 2000).

From the educational perspective, the typology of activity varies depending on the public involved in the research process: contributory projects, collaborative projects and co-created projects (CAISE Inquiry Group, 2009). In contributory projects, the public's mission is basically to contribute with data, observe and collect samples. Collaborative projects call for a greater implication; in those cases the public does not only contribute with data, but also help to redesign the projects, analyze data or discuss the results. The co-creation of science projects requires further compromise, as the participants are involved in most of the scientific process from helping to ask research questions and designing the study to interpreting and disseminating results.

The evaluation of informal science projects in terms of impact in education measures scientific awareness, knowledge and understanding, engagement or interest, scientific skills (study design, collection, data analysis, etc.), attitudes, behaviours or social and economic impact (CAISE Inquiry Group, 2009). In a contributory project such as The Birdhouse Network of the Cornell Laboratory of Ornithology, Brossart et al. maintain that the project had an impact in knowledge of the particular topic (bird biology), however it did no detect change of attitude toward science and the environment. In this case the participants had a previous positive posture, above normal, towards science and the environment(Brossard et al., 2005). Other studies, as in the case of the Citizen Sky project [3], noted a positive change in their attitudes towards science in projects that actively engage participants independently of the typology of contribution, as well as the importance of a social component, creating a community of knowledge and empowering participants (Price and Lee, 2013).

Collecting large quantities of data requires significant efforts in terms of recruitment, engagement and data quality, but also in striving to achieve learning goals. Therefore, the development of a project demands a careful planning in order to accomplish the project's protocols and, particularly concerning the formal learning perspective, provide educational material to support participant understanding in formal learning (Bonney et al., 2009). The study of the effectiveness to reach educational goals revolve around concepts like participation on scientific literacy (Crall et al., 2013; Bonney et al., 2009), knowledge of the topic(Bonney et al., 2009; Jordan et al., 2011) and changes in attitudes and behaviours (Crall et al., 2013; Jordan et al., 2011; Vitone et al., 2016). Jennett et al. (2016) develop a model to study motivation, learning and creativity. Regarding learning, they suggest that the learning occurs when the participants take part in tasks such as contributing, interacting with others, using external resources, using the project's documentation and sharing personal creations.

The studies developed in this thesis, especially in Chapter 2, shed light on educational outcomes in participatory experiments and crowdsourcing activities, fundamentally the change of attitude and science disposition that emerge of the participation in it. Involving the general public in scientific activities generates an indirect opportunity to learn in different scales like a particular topic of study, the scientific process or even the influence on attitude changes and behaviours. In participatory experiments, we have a direct contact with volunteers that enhance the possibility to exchange information and knowledge. However, in crowdsourcing projects, where the volunteers are widely dispersed across the globe and there is no face-to-face contact with

3. Citizen Sky. http://www.aavso.org/citizensky.

the participants, the platform and activity designs are a critical point in order to transmit the concepts and knowledge.

1.4 Data Analysis

Data has always been fundamental in scientific research. Research data are data collected, observed, or created for purposes of analysis to produce the original research results through simulations, interviews, observations, surveys, experiments, or even other resources.

Especially since the beginning of the 21st century, thanks to the digitalization of large processes, a large amount of data has been generated, captured, stored and analyzed in multiple areas. Disciplines such as computer science or software engineering have dealt with data management. At the beginning, the advances in data science were technical, improving techniques for managing and operating with data. Progressively, these advances bring data science closer to unrelated disciplines and governments, businesses, scientists, activists and even artists have understood quite quickly the potential of data, and their interest in getting acquainted with data increases year by year (Einav and Levin, 2014).

The expectations aroused by data, in particular by what is considered big data, have led to the emergence of disciplines that use data to study different phenomenons. A particular case is that of computational social sciences (Conte et al., 2012; Lazer et al., 2009), already mentioned above, which by definition is an interdisciplinary field mainly between social, computational and complex science. In this case, it is based on the use of social data generated in diverse digital platforms, especially (although not only), in digital social networks (Grimmer, 2015). These data sets formed by social data tell us about the behaviors of people in their daily lives, so they brings up issues regarding ethics, privacy, bias, fairness, and inclusion.

There are applications in several domains, describing common situations with social impact in our everyday life. The civic and political mobilization in May 2011 in Spain, known as 15M, was analysed using online networks data from Twitter. Researchers collected tweets that contained hashtags that referred to the protest during a period of a month (April 25 to May 25). The idea was to study patterns and find evidences of social influence and complex contagion. The study tested models of collective action empirically (González-Bailón et al., 2011).

Sometimes data from online networks is used provocatively. This is, for example, the case of Uber and the “Rides of Glory” post published (and then eliminated) in its data blog. They analysed data collected in their platform about rides, and from the analysis inferred patterns in the engagement in one-night stands. They concluded, for instance, that in nights such as the tax day or May 5th there was a higher probability of one-night stands than in Valentine’s day(Derrick Harris, 2012). From that example emerge some issues to consider when we are dealing with private data, but it also illustrates how companies use the data generated by us.

We currently have technological infrastructures that allow us to store and analyze these large amounts of data. In this last point, machine learning, a subfield of computer science, presents tools that help to analyse data through tasks such as prediction, classification or in the search for patterns.

In this work we focus on data at a scale much smaller than big data but larger than the data generated in typical behavioral experimentation. "Artisanal data", data sets that have been produced with the collection of data from controlled experiments allow, for instance, to hand-label or to detect biases while we are collecting. Part of the data analysis performed in the experiments is based on some simple machine learning techniques that help find patterns that we will introduce here, as well as some notes on data collection and privacy.

1.4.1 Machine learning

The idea behind machine learning is, as its name suggests, the ability of a machine to analyze data, extract information automatically and, finally, learn from it. Thanks to the great explosion in the availability of data and the reduction of the cost of computational processes, machine learning techniques have been adopted across many walks of life. The applications of machine learning go beyond computer science, where they were born and applied to fields such as computer vision, natural language processing, robot control, among others. Nowadays we can find applications across a range of multiple fields in science and industry (Egnor and Branson, 2016).

We can find well-known recommendation algorithms in online platforms of music, such as Spotify (Adam Pasick, 2015), or video streaming as Netflix (Libby Plummer, 2017). The Netflix Prize, a machine learning and data mining competition for movie rating prediction (Xavier Amatriain and Justin Basilico, 2012), is an example of the potential in the industry of machine learning and the importance for certain companies (as a note, The Netflix Prize was cancelled in 2010 due to privacy concerns). Social networks and IT companies have numerous applications, and along the same lines the game industry has made huge advances in classic games like chess (Shannon, 1950; Silver et al., 2017), or new ones like Super Mario World and Minecraft (Julie Muncy, 2016). Governments use machine learning in order to detect patterns of crime, criminal behavior or fraud (Rudin, 2013), and it is also used in economics for predicting stock markets (Patel et al., 2015). In natural science there are projects to identify endangered whales (Cornell Lab. Bioacoustics Research Program., 2018) or honeybees encounters' behaviours (Blut et al., 2017). Nonetheless, applications go beyond frontiers, reaching even art, where algorithms reproduce pictorial styles (Xavier Amatriain and Justin Basilico, 2012), and activism (Imran et al., 2016).

Particularly across empirical science, we can find applications from biology to social sciences. The latter, social sciences, is the most important application for us, and in which we will focus. Specifically in computational social science and the study of human behavior patterns.

Basically, there are three types of learning for machines. (a) Supervised learning, which is learning from a training set of labeled data provided to the machine. The classic examples are the spam classifiers, face recognizers or medical diagnosis systems. (b) Unsupervised learning, which looks for hidden structures or patterns in unlabeled data. (c) Reinforcement learning, a goal-oriented learning in which an artificial intelligent agent makes decision with the goal to maximize a reward. It is an intermediate approach between the supervised and unsupervised learning. Instead of training data that indicate the correct output for a given input, the algorithm receives an indication of whether the action is correct or not.

In this introduction, we will focus especially on some unsupervised learning techniques that have been used to look for hidden behavioural patterns in the data collected in the dyadic games experiment with emerging of phenotypes (Chapter 4) and in the collective-risk dilemma framed in climate change (Chapter 5).

1.4.2 The nature of the data

The data in the human behavior experiments is collected through the experimental platform that has been developed for this purpose (see Chapter 3). The data collection follows the protocols established by the legislation[4] in terms of security and privacy, and it is also approved by the ethics committees of the institutions that host the experiments. The experiment is performed following the citizen science practices described in the Section 1.3 and particularly in the work of Sagarra et al. (2016).

4. Organic law 15/1999 of December 13 about protection of personal data, (LOPD).

We collect a sequence of decisions made one after another during the human behaviour experiments described in depth in this work. Consequently, we obtain a set of actions that together establish behaviors or strategies in decision making. In general, the actions performed by users do not describe any behavior defined beforehand, therefore we are working with unlabeled data in terms of behaviour. Only in the case of the experiment in the field of mental health (see Chapter 6) we have the participants classified by groups, not regarding to the participants' behavior but to their mental condition. So, in this case. we could consider some analysis techniques with the data partially labeled.

Given the nature of the data and the questions that arise from them, and focused on discovering patterns of behavior, we will address unsupervised learning algorithms. Basically, we describe the bases of the algorithms that we have used in experiments in Chapters 4 and 5.

1.4.3 Unsupervised learning

Unsupervised learning refers to the problem of looking for hidden patterns and structures in unlabeled data. The data given to the learner is unlabeled, a priori there is no mechanism to evaluate a potential solution. Approaches to unsupervised learning are: clustering (k-means, hierarchical clustering, mixture models, etc.), anomaly detection, neural networks, and; latent variable models (signal separation, expectation-maximization algorithm, etc.).

The applications in which this learning techniques are used, alone or combined with supervised techniques, are very broad. Particularly, clustering techniques are present in recommendation systems, for instance in order to offer the same services to users of the same group (Xavier Amatriain and Justin Basilico, 2012); in the fight against crime, locating areas with high rates of crimes (Chen et al., 2004), or in education, identifying students with the same needs (Chen et al., 2004). Facebook uses computer vision algorithms to look for patterns in photographs (Soumith Chintala and Yann LeCun, 2016) and researchers in neuro-oncology use of clustering to detect tumors (Li et al., 2009). In the US Presidential Election of 2012, which Barack Obama won, exploratory data mining techniques, such as clustering, were used to group voters with certain interests (Blut et al., 2017).

One of the first well-known applications of clustering, in this case spatial, is the one that refers to John Snow's cholera map. In the middle of 19th century London, there was an outbreak of cholera which killed more than 700 people in less than a week in the Soho neighborhood. Dr. Snow, a popular doctor in the area, used to use maps in articles to illustrate his theories. After the cholera outbreak, he wrote down all the cases reported on a map.

As he was collecting data, some areas of the map stood out with a higher density of points,. These areas corresponded to Broad Street. He identified the source of the outbreak as the public water pump on Broad Street. In the affected area there were exceptions, non-infected people, mainly factory workers, who used private water pumps . The Broad Street pump closed down and the number of affected decreased abruptly. Snow used a map to illustrate the cluster of cholera cases around the pump and statistics to illustrate the link between the quality of the water source and cholera cases (Simon Rogers, 2013).

The data collected in the experiments described in Chapters 4 and 5 are sequences of actions that evolve over time. This data is complex enough so that it is not trivial to detect patterns of behavior in a simple way, so that automated algorithms are necessary. Basically, we have a vector with the decisions that each participant has made throughout the experiment performance. The objective is to find patterns in the actions of each participant that are repeated in a similar way for different participants, so we can group participants who behave in the same way, basically constructing clusters. A cluster is a collection of data which are similar between them, and dissimilar to data in other clusters.

Clustering is a task that basically consists in grouping objects. This task can be developed in multiple ways so there is a great diversity of models and algorithms. The paradigm of each model can make the construction of clusters vary greatly from one to another. We can find models that build the clusters based on the distance connectivity, like the hierarchical clustering. K-means uses a vector of means and the centroid,. Other cases use statistical distributions, density models, and so on.

The basic tools that are needed to make a cluster analysis are a proximity measure, a criterion function to do the clusters and an algorithm to compute. There are basically two criteria for the cluster formation: intra-cluster cohesion (how compact the cluster is), and the inter-cluster cohesion (how isolated are the clusters among them).

Among the multiple clustering techniques available we will focus on those that have been used in the development of this work, which are k-means and hierarchical clustering. We will also describe the methods applied to finding group formation and computing the robustness of each one.

K-means clustering

K-means is a clustering technique defined by MacQueen (1967), perhaps the most popular technique due to its simplicity and efficiency. Within the family of clustering techniques it is one of the partitional type (centroid).

The idea of k-means clustering is to divide our data into k groups. The output of the algorithm gives a label, a membership to a certain cluster defined by a centroid. Thus, the centroids of each cluster are adding the points closest to their cluster.

The k-means algorithm works as follows:

1. Define k centroids. Choose k random data points to be the initial centroids. There are algorithms to assign the initial centroid that ends up converging more efficiently.

2. Find closest centroid and update memberships. Assign each point to the nearest centroid, therefore to each cluster. The measure of the distance between points is a parameter (often Euclidean distances).

3. Re-compute the centroids based on the current cluster data points. The new position is calculated as the average position of all the data points in the cluster.

   (Repeat steps 2 and 3 until the convergence)

The convergence criterion is:

1. No (or minimum) re-assignments of data points to different clusters, or

2. no (or minimum) change of centroids, or

3. minimum decrease in the sum of squared error.

This cluster technique is sensitive to outliers, thus requiring additional pre-processing dealing with outliers (Zhou et al., 2009). In the same vein, it is sensitive to the seed (initial centroid), and there are techniques for the efficient management of this choice (Kang and Cho, 2009). The algorithm is only applicable if the mean is known, and with categorical variables there is a different version of the algorithm (Huang, 1997; Ralambondrainy and H., 1995). Finally, the number of clusters (k) has to be specified manually by the user. There are also a bunch of techniques to specify this parameter (Charrad et al., 2014; Pham et al., 2005; Tibshirani et al., 2001), but it is a fundamental criteria in order to build the correct number of cluster.

Hierarchical clustering

Hierarchical clustering is a cluster analysis method that builds a hierarchy of clusters. The strategy for its construction is based on two premises that divide the hierarchical clusters into two types:

- Agglomerative: this is a "bottom-up" approach. Initially each data point is in its own cluster, and the algorithm evolves merging points as one moves up the hierarchy.
- Divisive: this is a "top-down" approach. Initially the data points are in a unique cluster, and the algorithm evolves the cluster as one moves down the hierarchy.

The agglomerative hierarchical algorithm works as follows[5]:

1. Start with N clusters. One cluster for each data point.
2. Merge the two clusters that are closest to each other.
3. Re-compute the distances between the clusters.

   (Repeat steps 2 and 3 until get one cluster of N points)
4. In the dendrogram (the tree representation of the clusters) choose the number of cluster (k).

This algorithm needs dissimilarity measures to decide which clusters merge. On the one hand, there is the metric that measures the distances between pairs of data points, and on the other hand there is the linkage criterion that specifies the dissimilarity of the sets of data points according to pairwise distances of the observations in the set of data points. The metric is the distance between two points, and the distance modifies the shape of the clusters. The points may be close or far depending on the distance. Some common distance in hierarchical clusters are: Euclidean, squared Euclidean, Manhattan or Mahalanobis. The linkage method, which can be complete, single, center or Ward's, determines the distances between a set of points.

Evaluation and robustness

The use of unlabelled data makes it difficult to reliably evaluate the clusters that have been created because there is no "ground truth" classification. There are some indexes (Charrad et al., 2014) that help us find the algorithm that works best with our data or the optimal cluster number in our dataset given a particular algorithm.

Let's take as an example to illustrate this one of the indices that have been used in the experiment detailed in Chapter 4. This is the case of the Davies-Bouldin index (Davies and Bouldin, 1979), an internal evaluation scheme that validates the clusters with the features of the dataset. Basically, it calculates an index that defines the ratio of the within cluster scatter to the between cluster separation. In other words, the number of clusters is such that it presents the minimum dispersion within each cluster, and the maximum distance between all pairs of clusters. The best clustering scheme minimizes the Davies-Bouldin index, hence is a good measure of how many clusters actually exist in a given dataset and therefore the data could be ideally classified. Its use is very common for algorithms such as the k-means or hierarchical, in which it is necessary to identify the number of clusters.

However, these indices may vary if the algorithms used are susceptible to seeds; this is the case of the k-means. There are some techniques that allow indices to be calculated in a way

5. The agglomerative method has been used in the analysis (hclust R package).

that is more robust, independently of the seed. In the study presented in Chapter 4 we ran the algorithm 200 times on our data with different seeds for the algorithm in every run. Once the 200 indexes were computed, we used the average and the standard deviation value of the DB-index to choose the k.

To analyze the robustness of our clustering analysis against data perturbations, we could run the algorithm with a subset of the original data N times and look for the threshold from which the results do not show significant differences with the whole data set. Concretely, we run the algorithm 200 times excluding a given number of participants' actions randomly chosen. By doing so, we allow the minimum number of samples that allow a number that does not present a significant difference with the full data set.

The participants that compose the clusters can change each time the algorithm is executed, so there is a need to compare the consistency of the composition of the clusters each time an algorithm is executed. One measure of cluster similarity that we used in the study is normalized mutual information score (MacKay, 2005). The Mutual Information is a measure of the similarity between two clustering systems of the same data into disjoint subsets, and it is given by the relative entropy between the joint distribution and the product distribution, 0 (no mutual information) and 1 (perfect correlation).

Consensus clustering

As previously mentioned, unsupervised techniques have certain shortcomings, especially in the interpretation of results, which must be overcome with a thorough testing of the results. From the get-go, not knowing the number of groups that best suits our data and not having an external verification system, it is difficult to ensure that the results of the analysis are correct.

We have referred to some techniques to overcome these limitations and check the robustness. The Consensus Clustering (Monti et al., 2003) is a method that gives evidence about the number and membership of the clusters that best fit to the data. It provides the result of consensus across multiple runs of a clustering algorithm, allowing the implementation of several algorithms (k-means, hierarchical, etc.) determining the number of optimal clusters in our data and providing stability values of the clusters that have been formed. Basically, it calculates two important stability parameters: Item-consensus (IC) and Cluster-consensus (CC). IC is the average consensus value between an item and members of a consensus cluster, so that there are multiple IC values for an item at a k corresponding to the k clusters (Wilkerson and Hayes, 2010) and CC the mean of all pairwise consensus values between a cluster's members, so there is a CC value of clusters at each k (Wilkerson and Hayes, 2010). Furthermore, both consensus rates are the metrics used for estimating the optimal cluster number (k).

Consensus Clustering helps identify the number of groups and the membership in them. It is widely used in biomedicine (Verhaak et al., 2010), as well as in other disciplines such as complex networks (Lancichinetti and Fortunato, 2012). However, it also has its own limitations (enbabaolu et al., 2015) that can be minimized with the use of different heuristics (Lancichinetti and Fortunato, 2012) or other implementations (Krieger and Green, 1999; Lock and Dunson, 2013; Burgess et al., 2016) that have been developed to solve specific problems or the way in which the optimal number of clusters and their robustness is discovered.

In the work presented in Chapter 5, the R ConsensusClusterPlus (Wilkerson and Hayes, 2010) package was used to perform the clustering analysis.

Part II

Participatory Platforms

Chapter 2

Natural Patterns: A Participatory Experience in the Wild

SUMMARY – Citizen science provides a tool to rebuild public confidence in science and reduce the gap between science and society. We present Natural Patterns, a platform that introduces the scientific method to new audiences, providing game-based activities that promote the observation and collection of patterns, the creation and discussion of hypothesis and the participation in challenges. We introduce an entirely new conceptual approach to designing citizen science systems, which serve as a catalyst to promote scientific reasoning and to increase the engagement in science. We describe a set of design principles to boost the motivation of participants in scientific activities and to promote the understanding of science through citizen science projects.

2.1 Introduction

Citizen science (CS) practices (Gura, 2013; Hand, 2010) try to reduce the gap between science and society (Leshner, 2015), promoting the participation in scientific research and changing attitudes towards science (Bonney et al., 2015). CS, and particularly crowdsourcing based projects, allow non-expert citizens who, in principle, do not have any role in science, to take an active part in real scientific research. Involving people with no formal training in science and engaging them in scientific activities could encourage learning, promote scientific knowledge (Bonney et al., 2014; Cooper et al., 2014) and create a more positive feeling towards science (Bonney et al., 2015).

Most of citizen science projects do not draw on a broad base of participants. Particularly, projects based on crowd participation are successful thanks to a small subset of dedicated participants interested in the topic of study (Sauermann and Franzoni, 2015). Besides, most of the projects can be categorised as contributory projects (Roy et al., 2012), based on the scientific process to which participants contribute (Haklay, 2013; Roy et al., 2012; Bonney et al., 2009). This leads to CS projects that not only fail in captivating and retaining the participants, but also do not offer participants the opportunity to achieve alternative goals in learning, science disposition or simply enjoying the participation, to mention just a few, because most of the projects are mainly research-oriented, leaving the participants motivations aside (Raddick et al., 1988).

Instead, CS activities need to be designed beyond data collection and need to engage citizens in the entire scientific process (Bonney et al., 2009; Cohn, 2008). They have to include extra elements in their design: first, they have to add alternative goals and features that motivate the participation of the general public (Jennett et al., 2016), and second, they have to create

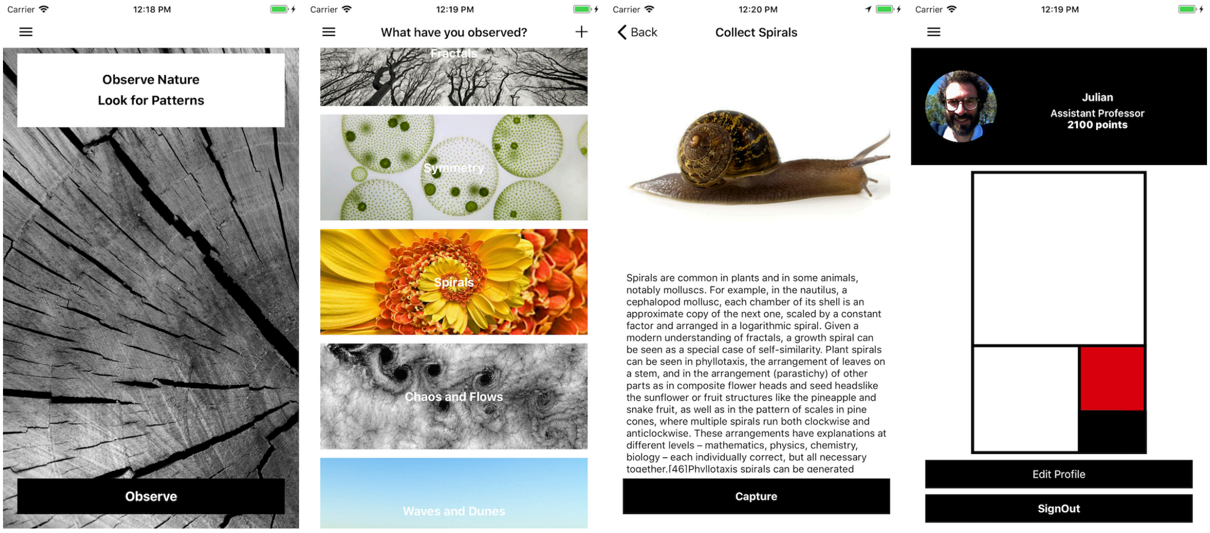

Figure 2.1: Natural Patterns application. Screenshots of Natural Patterns main interfaces focused on the observation of nature, the classification of patterns and the capture of samples.

participatory environments that promote long-term science engagement (Qaurooni et al., 2016). Finding an equilibrium between the motivations of the researchers and the public is the key to encourage new audiences to participate and to improve the participant's overall experience while maintaining the scientific rigour of the project.

Our conceptual approach is to design CS experiences that are primarily aimed at engaging a broad base of people in citizen science activities and that do promote the engagement and comprehension of the scientific process. While data is already being collected through CS experiences that embody this approach, our focus is not primarily on using that data for scientists per se, but instead, on using the data collection process and the scientific questions it raises (e.g. through thinking, through hypotheses) to simultaneously engage people more fully in the scientific process and to promote their interest in doing CS. Longer term, such efforts can potentially increase the public's interest in science, and build a broader base of science-interested participants who may be willing to engage with other CS experiences or other activities that contribute to science.

In this paper we address the challenge of designing and implementing a platform called Natural Patterns that proposes a participatory experience in which the participants, following the steps of the scientific method, have to look for patterns in nature. The system does not only engage users to collect patterns, it encourages them to participate in multiples stages of the scientific research process: observe the natural world, classify samples, form hypothesis and provide evidences for or against particular hypothesis.

We design the system so that the topic, patterns, tries to highlight the implications of science particularly in nature but also in different contexts of our day-to-day life in a broader sense (architecture, art, objects, etc.). The flow of the platform illustrates the main steps of the scientific method, and the activities and microtasks are designed to interact with nature and with other participants. The core mechanisms of the platform try to promote participation and engage the users by motivating them (e.g. offering rewards when goals are achieved).

Natural Patterns provides a platform to study how CS can have a positive impact in science disposition in the long term and how to engage the general public in science activities out of the lab. Here, we want to know if the main design features of Natural Patterns help to create strong relations among citizens, science and nature. Because, in essence, by connecting us to nature and promoting environmental observation we can understand human impact in nature and the importance of science to find solutions to social problems (Dickinson et al., 2012).

To evaluate Natural Patterns we performed a study in which each participant employed the platform during 10-days. A total of 17 participants took part in it. During the study, we collected data about their usage of the application, focusing primarily on the interactions with the main application features previously described. Additionally, we also asked the participants to answer a survey about science disposition and the scientific method prior to the first usage of the platform (n=17), and a post experience survey (n=8) in which we wanted to find out how the experience had impacted in the participants' science disposition, and how the platform features represent the scientific method, motivate the participants and promote interaction with their surroundings, specially with nature. Besides that, we completed the study with specific questions about the general user experience.

The results reported in this paper indicate that by using the platform the participants interacted with nature more often than usual. Participants also report that most of the scientific research steps are well-represented, especially observation, collection, hypothesis and conclusions.

This paper makes the following contributions:

- Presents a platform that proposes activities focused on applying the scientific method to the topic of finding natural patterns, engaging the public for the benefit of science and promoting science disposition in the long-term.
- Validates the design features of activities created to improve the CS experience beyond scientific outcomes – critical thinking, engagement, rewarding, competition, observation, learning or scientific discussion.
- Evaluates if the design features included in the platform replicate the main steps of the scientific method, and sheds light into which features need to be redesigned to improve the whole scientific experience.

2.2 Related Work

2.2.1 Beyond the scientific outcome

Some of the main goals that citizen science projects pursue go beyond the production of scientific outcomes. They want to increase the scientific thinking of the participants, their understanding of science, and they also want to have an impact on community problems, increasing the relation between environment and society. Trumbull et al. (2000) highlights the importance of studying participation and discusses how participation should strive to address misunderstandings of science and how citizen science projects could create an ideal forum for sharing thoughts.

Participation reinforces the participant's knowledge (Bonney et al., 2015; Brossard et al., 2005); however, the impact in the participant's attitudes toward science or the understanding of the scientific research process is not clear (Price and Lee, 2013; Brossard et al., 2005). Projects that try to achieve ambitious targets such as social change or to improve science-society interactions need to pay special attention to aspects related to the project's design, to the engagement of new audiences, to measuring outcomes and crossing research boundaries (Bonney et al., 2015). Precisely, these aspects are the nucleus of this investigation.

Crowdsourcing applications have been introduced in a wide range of fields. We find studies in environmental science (Massung et al., 2013), noise mapping (D'Hondt et al., 2013), emergency management (Gao et al., 2011) in several natural disasters like Haiti (Norheim-Hagtun and

Meier, 2010) or Fukushima (Brown et al., 2016), and currently crowdsourcing extends to the physical to solve societal challenges by taking advantage of mobile crowds seeking location-based tasks (Harburg et al., 2015). Part of the studies related to those projects have been focused on promoting social responsibilities or behavioural changes.

The studies of mobile crowdsourcing applications that address the outcomes for the participants have increased recently. Participation in those projects cover a wide range of non-scientific goals (educational, entertainment, change of habits or attitudes, etc.), and it does so at the same level that the reward for the scientist, the scientific outcome. The GLOBE program (NASA, 2017b) is a scientific and educational program designed to participate in data collection, and the scientific research process focuses on the Earth's system and global environment, creating a vast community to explore nature. Recently NASA launched Globe Observer (NASA, 2017a) and an international citizen science initiative, Clouds, that allows observations of clouds in order to study our changing Earth system. Likewise, another related application is iSeeChange(NASA, 2017c), a community that combines citizen observation, weather and satellite data to study climate change. Our platform supports the creation of a substantially larger number of activities related with different disciplines from Earth science to architecture (Museum of Science and Industry, 2017; National Geographic, 2017; Brandon Keim, 2017; Jess McNally, 2017; Watercutter, 2017). Nevertheless, our approach to topics such as those mentioned above, clouds or climate change, is done by creating activities under the umbrella of patterns in nature, the guiding principle during the experience.

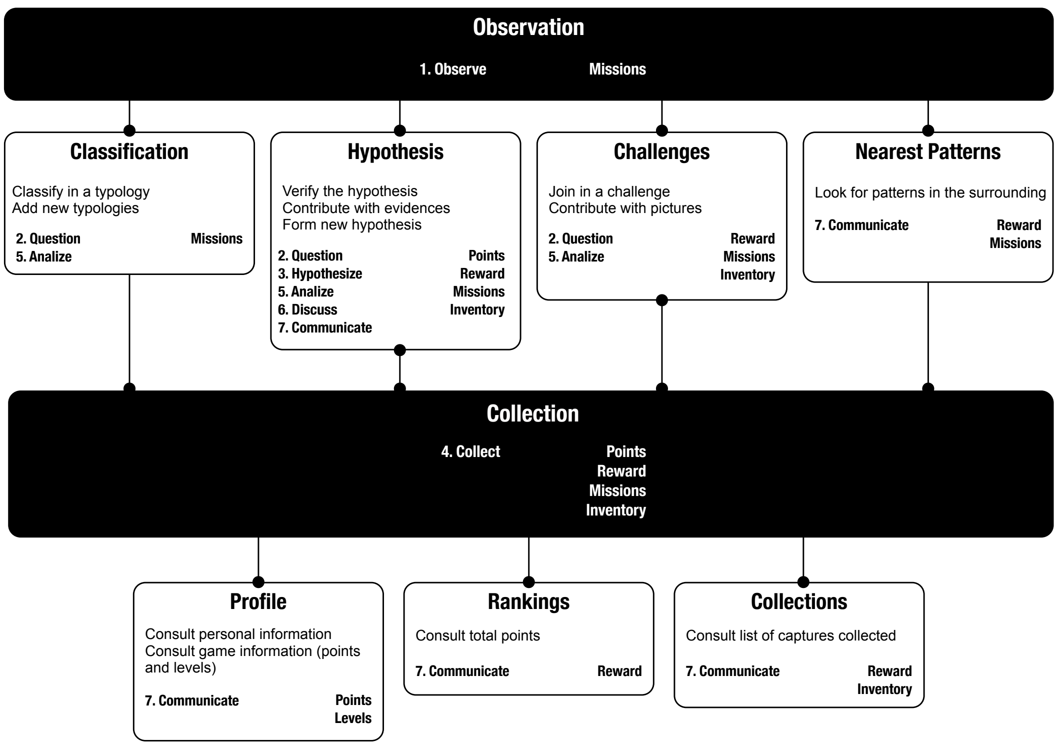

Figure 2.2: Natural Patterns block diagram. With the main interfaces and the relationship between them, with special relevance for the core interfaces of *Observation* and *Collection*. Each interface is accompanied by the main actions, the steps of the scientific research process and the game components.

2.2.2 Participation and motivations

Understanding the motivational factors that affect the participants, as well as the factors that drive people to take part of a community, is basic to understand what parts of the design need to be considered with special attention in order to facilitate participation and engagement. But, at the same time, those motivations are complex and changing (Rotman et al., 2012). There is a large body of work in motivation theory. Deci and Ryan (Ryan and Deci, 2000; Deci and Ryan, 1985) proposed a model of motivational factors classified into either intrinsic (doing something for its inherent satisfaction) or extrinsic motivations (doing something to get some separable outcome).

In citizen science crowdsourcing there are some studies of participation that distinguish both intrinsic and extrinsic motivations of their participants. With regard to the intrinsic motivations, some of them are: the aspiration of discovery, fun, helping science or contributing in a community (Curtis, 2015; Bowser et al., 2013; Rotman et al., 2012; Nov et al., 2011; Raddick et al., 2010). On the other hand, personal interest, acknowledgment, socialization, community sense or learning (Curtis, 2015; Bowser et al., 2013; Rotman et al., 2012; Nov et al., 2011; Raddick et al., 2010) are some of the extrinsic motivations found among the participants of citizen science projects. We laid out the experience emphasizing the motivational factors thought design elements. Afterwards, we studied the motivations of participation and how design features affect them, combining intrinsic and extrinsic motivations.

Along these lines and with the objective of opening the participation to new targets and increasing motivational components, we designed the experience as a game. Games in participatory science are built to cross boundaries and create interactions among different areas (Shapiro and Squire, 2011). The effects of gamified experiences in participation are well-documented in literature. Morschheuser et al. (Morschheuser et al., 2016b) review crowdsourcing applications arguing that participation increases and the gamification mechanisms vary according to the activities, in fact the most common gamification affordances are points/scores, ranking and badges. The impact is diverse among projects between behavioural and psychological (fun, attitude and motivation) outcomes. Particularly in citizen science, gamification is applied to popular projects like FoldIt (Curtis, 2015) or Galaxy Zoo (Raddick et al., 2010). Game-like projects introduce new motivations not present in citizen science volunteers (e.g competition) and increase others (e.g: fun)(Bowser et al., 2013). In our study we measured the impact of gamification affordances in the user experience and the perception of game.

Natural Patterns creates a link between nature, science and the general public. This link is created through a crowdsourcing platform in which the participants collectively contribute with samples. The platform is presented as a game with microtasks, which are designed as scientific activities. To design an efficient system it is necessary to create activities that motivate the general public to participate in citizen science crowdsourcing.

2.3 Design Goals

Designing CS crowdsourcing applications implies facing design decisions based on heterogeneous but complementary objectives. The high-level goals that drive the design of Natural Patterns are the following:

- Promote interaction between participants and nature. The platform forces participants to focus on nature more than in the application itself, with features designed to impose the relationship with the environment (e.g. look for a specific pattern, take and share pictures).

- Design scientific activities that follow the principal steps of the scientific method (observe, capture, collect, analyze and discuss) to increase scientific and critical thinking.
- Enhance social interaction with other participants, fostering the scientific discussion. We allow the interaction with the data collected by participants in different ways.
- Create an experience that promotes public understanding of science and learning by doing scientific activities in the wild.
- Increase the interest of the public in science-related activities and improve the long-term engagement of the participants.

2.4 System Description

2.4.1 User interface

We developed a crowdsourcing mobile application, Natural Patterns, by which the participants study patterns following the main steps of the scientific method. In this context, Natural Patterns proposes an approach that we could describe in three layers that try to bring about the design goals. The first layer is formed by a set of activities with microtasks that every participant completes by using the design features of the application. In the second layer, the participants evolve in the understanding of the scientific research process by combining the contributions resulting from each micro-task. Finally the third layer introduces the game mechanism over the whole research process to increase long-term user engagement in the application.

The workflow from the user point of view starts with the creation of an user profile with basic information (username, password and picture). From the profile, it is possible to follow the activity on the application through two parameters: number of points and academic rank. A visual score pattern also gives feedback about the progress. Accessing from profile, there is a ranking which shows a list of users sorted by number of points derived from their contributions during the performance.

The scientific method has been the subject of intense debate throughout the history of science from Aristotle's Organon (4th century B.C), through the Baconian method (1620), to nowadays. Basically, in our context, the scientific method is used to discover cause and effect relationships by asking questions, capturing and examining samples, and inferring logical answers in an iterative process. A assortment of micro-task has been implemented to create the scientific method layer. In other words, the experience's flow basis is: observe, capture, collect, analyze and discuss. The interfaces allow the following interactions:

Captures. This interface is the combination of ***Observe*** and ***Collect***, and it presents a path connecting the application with nature, intending to extrapolate the patterns presented in the device and trying to find replicas in the natural world. We propose examples, with pictures and explanations, of basic and well-known patterns that it is possible to find in nature (fractals, symmetry, spots and stripes, spirals, chaos and flows, waves and dunes, bubbles, tessellations, and finally, cracks). Besides, new patterns, not contemplated in the list, can be added easily just typing a pattern's name and uploading a picture. Once the pattern is uploaded, it appears in the list of patterns along the author's name. It is the core of user interaction, the main task is to collect (taking pictures) and classify samples, share, analyze and compare them with the rest of participants. Once a participant observes a pattern in his/her surroundings, he/she takes a picture and adds it to the collections gallery, together with the patterns of the rest of participants. It can be captured from the app or be uploaded from the photo library. The usual situation is, first, to classify the sample in the correct taxonomy and, subsequently, to capture the samples using the camera. The picture has a metadata associated, mainly the location, allowing to detect samples in the real context and to create new features associated with each one.

Challenges. Periodically, a challenge is proposed to the community with the objective to deepen a particular subject. In the first place, we present a list of challenges open to contributions during a limited period of time. This period is variable from a day to a year in function of the objective of the study. In the case of challenges, when they are activated the participant can join and contribute with samples, basically taking a picture that responds to the call. Once the challenge is finished, the participants receive feedback with a brief study about the contributions, and the participant who contributed with the most interesting data receives a reward. The challenge has different scales of time and applications. Some examples of challenges are: (a) "Solar Eclipse: during a solar eclipse day take a picture of the sun every five minutes", by which we obtain the evolution of the eclipse in time and space through recollecting all the geolocated samples; (b) "Pattern Evolution: Choose a pattern near you and take a picture every week.", which allows us to observe the evolution of a pattern during a large period of time, for example along different seasons; (c) "Clouds: Collect samples of different types of clouds." This section opens up a wide variety of possibilities for creating activities, even relating the nature with other context (architecture, design of objects, and so on).

Hypothesis. The interface responsible for displaying hypothesis is designed to explain the idea that underlies behind the hypothesis concept, add evidences that support (or not support) each statement, and understand what hypothesis have been proved or unproved at a glance. The players are able to create new hypotheses that can be supported (or not) with the data collected. Once users access a hypothesis, the application asks for their opinion about the statement (agree or disagree) and to provide evidences that support the choice.

Patterns near you. Look for patterns in locations near you, in places where other participants have been collecting data. This is done through a list of the places near the user (maximum distance of 1000 meters) and a map with the points where the data has been collected.

Collections. Per each participant we show all the data captured, classified by date and typology of patterns.

Among all the interactions and microtasks there are a variety of levels of action, from lowest to highest. *Collections* requires passive activity, just information queries. *Patterns Near You* acts similarly but, nonetheless, it adds an action component, the idea to look for patterns in the vicinity. *Observe* promotes the action of observation in the wild, an interconnection between the device and the natural world. *Collect* takes advantage of *Observation* focusing on the device introducing the classification of data and the action of taking pictures. Finally, the highest levels of action are present in *Hypothesis* and *Challenges*, a mixture of observation, collection and analysis.

2.4.2 Game design

Following the *Gamification Model Canvas*, a tool to design a gamified experiences based on Hunicke et al. (2004), as an inspiration, we decide to create a complete playable experience. The platform in which we developed the experience is a mobile phone with iOS but the activity also takes place in real scenarios, so the physical context is actually an important feature of the game. We create a multilateral competition which allows us to promote social interaction among players. The dynamics of the game is created by mechanisms like observing the surroundings, capturing samples (50p), defining new patterns (100p), proposing new hypotheses (200p), providing (counter) evidences (50p) to support hypotheses (25p) or dealing with challenges (50p), among others. Therefore, in terms of game dynamics, the exploration of the physical space

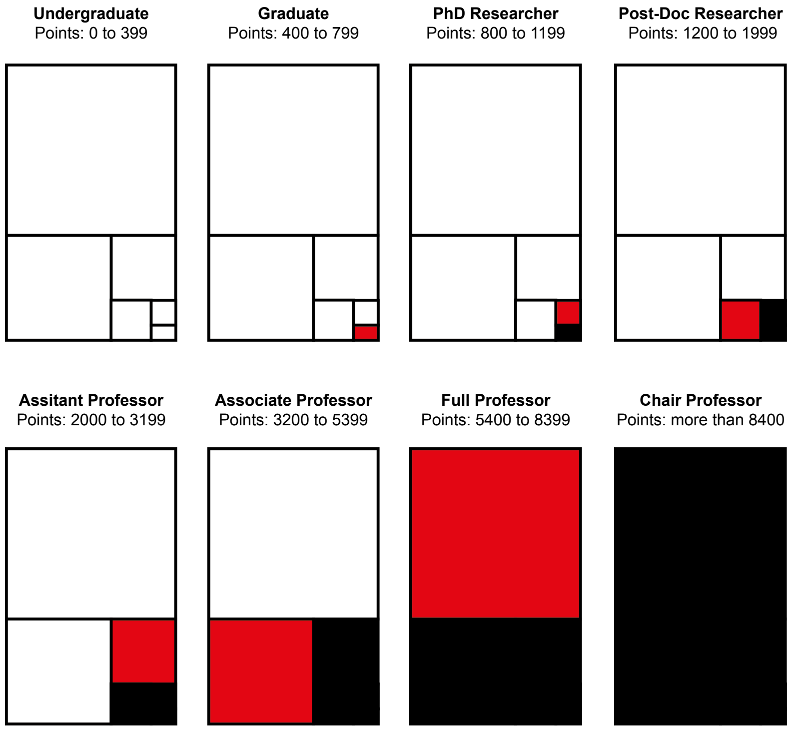

Figure 2.3: Fibonacci academic progression. Levels and academic position of the game following the Fibonacci sequence.

incentives discovery, the achievement of a challenge and objectives makes the participant's status grow and incentives competition, and looking for patterns near you stimulates fellowship and socialization. The main game affordances are points, reward, levels, missions and inventory. Finally, we design the game experience to evoke emotional responses to the player though aesthetics: challenge, fellowship, exploration and discovery (Fig.2.3).

2.4.3 Technical details

Natural Patterns has to accomplish the design goals by means of the interactions and features previously explained. The main challenges have been solved as described below.

Observation has the objective to promote interaction between participants and nature, trying to strengthen the observation of natural details in the wild, and to take the eyes off the device screen. We create a bidirectional interaction where nature is the central object. In *Observation*, the interaction is from nature to device. Nevertheless, in *Challenges*, *Hypothesis* or even *Patterns Near You* the information shown in the device (i.e: a challenge about clouds, a hypothesis about symmetries or a pattern near our work) create interaction both ways, from and to the device.

All features are connected. The combination of features create a layer that represents the scientific method, so the application flow simulates a first approximation to the real scientific research process. At least, it simulates the basic steps: observe, hypothesize, collect and analyze data, discuss and communicate.

Challenges, *Hypothesis* and *Patterns Near You* are social activities, the idea to create a community that shares and discusses. Those features promote actions in and out the device and ask for contributions that make sense, especially, in collaboration with others, not individually. As an example, the challenges create a competition with others, comparing samples; the hypothesis are discussed and evaluated collectively to be accepted or rejected, and, finally, the nearest patterns exist if other participants collect patterns in their surroundings.

Above the complete experience, we design a game layer that introduces some components of feedback and reward. Users can find their personal information in *Profile*, and the information with respect to the others in *Ranking*. Each contribution has a score (e.g. capturing a samples are 50 points), and, as it increases, so does the participant's academic rank as a researcher (e.g. with 1600 points, you are a postdoc researcher). The rank increases following the pattern of the *Fibonacci* sequence.

To improve participation and intensify exploration and learning in participants, we create periodic activities like new ideas to collect patterns, so the application is dynamic in its use. There is a trend to always capture the same well-known patterns (fractals in trees or spirals in flowers). However, a lot of patterns emerge, some of them during short periods of time. The feature *Challenges* promotes the discovery of new patterns and even the study of the evolution of patterns, enabling the creation of large and vast collections of natural phenomena (clouds formation, changes of patterns during seasons, and so...)

2.5 Study

We conducted a 10-days study with the objectives of (1) evaluating whether the design features of Natural Patterns represent the scientific research process, and (2) examining whether the design features of Natural Patterns affect people's attitudes and dispositions toward science.

In order to do that, we recruited 17 people who had an iPhone 5S or above and iOS 10+ via university mailing lists and social networks. 47.06% of the participants were female and the average age was 31.94 (SD: 7.27).

The study was performed in three phases: (1) We asked participants to complete a pre-survey (n=17) that collects information about science disposition and understanding of the scientific method before using the application. (2) Then, those who completed it in the agreed time were instructed to download the application (n=15) and use it (n=12) for a period of 10 days. Meanwhile, we were collecting data about their participation. (3) Subsequently, the participants who performed the minimum tasks required were asked to answer a post-survey (n=8) questioning about design features, participation experience, as well as queries about science disposition and the scientific method, like in the pre-survey.

We distributed some simple instructions among the participants together with the main functionalities of the application. Besides the pre and post survey, the participants had to carry out a set of daily actions so that the study of the experience was valid. Participants who completed all phases of the study correctly received a 15 euros gift card as compensation for their time.

During the experiment, the participants (n=12) captured a total of 115 samples and interacted with the application's main features 515 times. Figure 2.4 illustrates the number of active participants and the number of samples collected every day.

Likewise, the samples are collected from the microtasks challenge, hypothesis, and from the general feature Collect/Capture – 59 samples (51.3%) where collected using this function. The distribution of the collected data is also represented in Fig.2.4.

2.5.1 Natural Patterns features

The features were designed and oriented to very clear objectives. Natural Patterns had to promote interaction with nature, to motivate the participants intrinsically and extrinsically and to enhance critical thinking, among many other things. Below, we present the results of the evaluation of the participants in the study (n = 8). Firstly, the evaluation of all sections of the

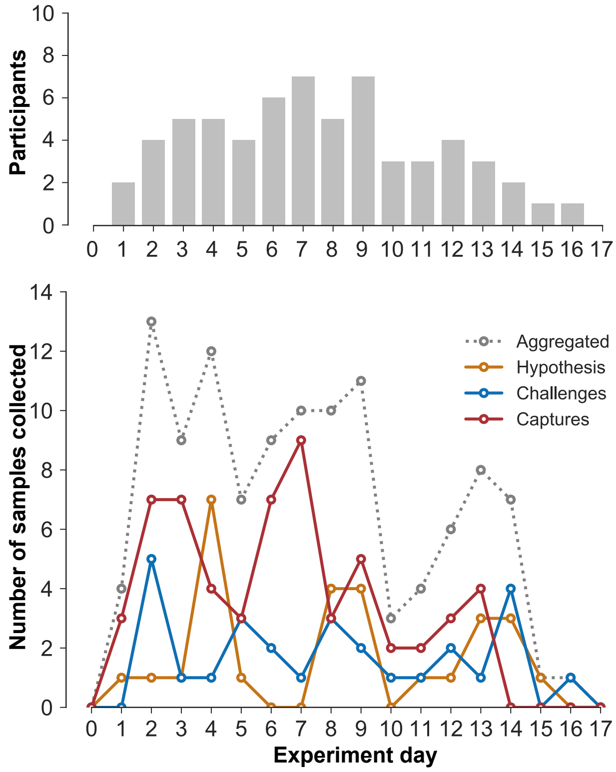

Figure 2.4: Activity during the study of Natural Patterns. (Top) Number of participants collecting patterns and (Bottom) distribution of captures.

application, and secondly, the assessment of the features of the main microtasks implemented: *Captures*, *Hypothesis* and *Challenges*.

Natural Patterns activities and microtasks

The participants rated each section of the application based on its utility. All the sections are rated over 2.5 (out of 4) except the section *Nearest Patterns*. The core sections – *Observe/Collect*, *Challenge* and *Hypothesis* – were highly rated.

We asked the participants for improvements in their favourite sections. *Observe/Collect* is one of the favourite sections. Participants' recommendations are addressed at the improvement of the explanations (*"I will add more patterns and different photos of examples and a shorter description."*) and at the ranking of collections (*"It could be cool if others could rank my collections."*). In *Challenges*, the suggestions are related with socialization (*"I would prefer that we can see all the other people upload"* or *"To compete with the other people"*) and playability (*"I would add scores and goals."*). Finally in the *Hypothesis* section, participants recommended socialization as well (*"[...]interact between researchers[...]"*) and engagement (*"[...]it should be more participative and engaging. It should enhance the discussion about the hypothesis."*)

We also asked for sections susceptible to be deleted or deeply improved. The most noted is *Nearest Patterns*, as some people did not understand the goal, or the section did not work (*"I have not understood the function of the nearest patterns"* or *"I couldn't use it."*). As an interesting point to improve collection, a participant asked for more community activities and rewards, and other participant needed basic functionalities to modify previous actions (e.g. adding more information a posteriori).

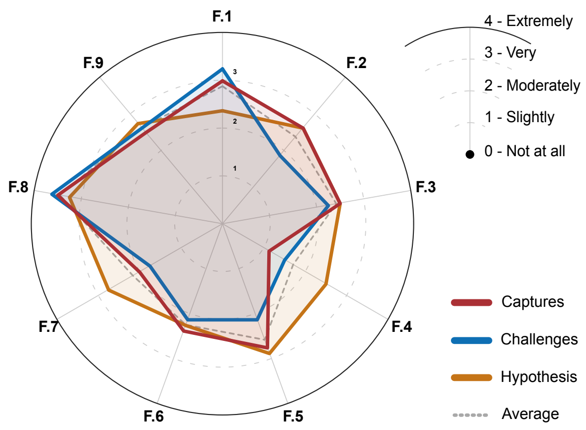

Figure 2.5: Evaluation of the Natural Patterns features by participants. The parameters are: F.1 Nature interaction, F.2 Nature understanding, F.3 Rewarding, F.4 Socialization, F.5 Critical thinking, F.6 Competition, F.7 Discussion, F.8 Observation and F.9 Engagement.

Captures

The interfaces of *Observation/Collection* are grouped as *Captures*, which has features that increase the observation of nature and the observation of surroundings in general. The other features are well represented, except the one that incentives socialization. The participants rate their experience interacting with this interface 3.12 (SE: 0.29). Some representative comments of the participants are: *"It was inspiring and funny"*; *"I enjoyed collecting patterns however tagging was a little complicated because I wasn't sure if I did well or not."*; *"It was interesting but without motivation in order to get a goal."*; *"It's been really fun and I've learned so much information about natural patterns that I didn't know. [...]" or "I have a lot of fun, but maybe more feedback and comments between participants will improve the experience"*. There are other interesting suggestions such as: evaluate or/and comment the photos of others, classify the pictures in multi-categories or the possibility to edit the photos once uploaded.

Challenges

Challenges was rated lower than the other interfaces, 2.62 (SE: 0.32). Observation of nature and surroundings are the main actions that the interface promotes. Also, it shows a low performance in socialization, understanding of patterns and discussion. This results reflect the need to improve the design of the interface because the idea under Challenges is to improve socialization and competition. In this case, the comments ask for improvements in the interface: *"There were some challenges I didn't understand so I didn't enjoy it or learning anything."*, *"It was nice to follow a challenge as a motivation"* or *"It's a good idea, but feels that need something else to engage the participant"*. There are suggestions about the reward and an increase of interaction with the others users: *"They are challenges, they should be rewarded somehow [...] A way to directly compete with other users should be implemented. I would like to challenge other users and see which is my overall performance compared with the others."*

Hypothesis

The experience interacting with the *Hypothesis* interface was rated 3.00 (SE:0.38). This interface specially promotes critical thinking and observation, as well as the other actions without exception. The prominent comments in this case are: *"It is interesting to learn new patterns about the nature thanks to the cooperation of other people"*, *"The most interesting aspect of hypothesis are the discussion the app doesn't provide any interface to do it."* or *"It's the best part - but need to improve the relationship between answers and participants."*

2.5.2 Scientific Method

The participants (n=17) were simply asked if they knew what it consisted of and to offer a simple definition of it. Most participants 58.82% considered that they know what the scientific method consists of, and 17.65% of the participants consider that they have some notions of it.

From the implementation of the Baconian method (1620), which underpins the basis of the scientific method in modern science, until now, the scientific method has been defined on countless occasions. Due the vast number of scientific disciplines, it is not easy to whittle down a unique definition. We could say that the scientific method is a body of techniques that characterize nature. Basically, the main concepts that define the scientific method are systematic observation, measurement and experiment, as well as the formulation, testing, and modification of hypotheses.

Most participants answered this question positively and 64.71% of the participants gave a definition close to the correct one. Al least, they introduced the main concepts of the scientific method: *"[...] fixing an hypothesis, testing the hypothesis with experiments in real life and checking from the results whether the hypothesis is right or wrong [...]"* or *"[...] Observation - Question - Hypothesis - Test - (Does it work? Draw Conclusions else GOTO Hypothesis) [...]"* just to give some examples. Nevertheless, two participants answered that they knew what the scientific method is however the definitions given do not reflect the idea of science inquiry as *"Ask, observe and apply."* or *"[...] method that try to prove all that happens surrounding us demonstrating the facts."*.

Natural Patterns as an analogy of scientific method

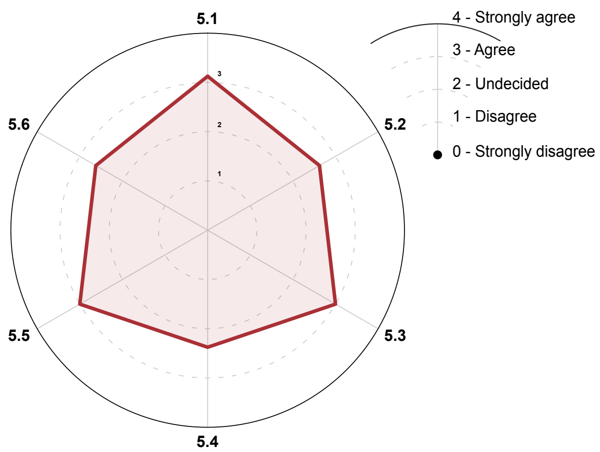

Figure 2.6: Study of Natural Patterns' features. Measurements of Natural Patterns' impact in the understanding of the scientific method. The parameters are: 5.1. Evidences relevance; 5.2. Collection procedures; 5.3. Samples importance; 5.4. Results summary; 5.5. Draw conclusions; and 5.6. Results communication.

Natural Patterns is designed with the goal to stimulate the understanding of the scientific research process and to introduce it into our day-to-day life. For this reason, the microtasks have components that represent the main steps of the scientific method. The participants valued how Natural Patterns, as a whole, represents and helps to better understand the multiple steps of the scientific method.

As Fig. 2.6 illustrates, Natural Patterns helps to understand what evidences are relevant to argue (or counter-argue) and the validity (or not) of a particular hypothesis. It helps to discern between the samples that are useful to solve an issue and those that are not. In the same way, it also helps to draw conclusions based on all the accumulated evidence. However, Natural Patterns does not have such a high impact in enhancing the knowledge of different methods of sample collection. It has a relative significance in the processes to summarize and integrate the results and, therefore, in the effective communication of results and conclusions.

2.5.3 Science disposition

To measure the scientific disposition, we will focus on two crucial aspects: science involvement and science perception. The participants initially (n=17) answered four questions about their involvement in science. 58.8% of participants do not take part in scientific activities professionally but, in the same percentage, they have done so in a non-professional ways (e.g. citizen science, crowdsourcing, etc.). On the same line, 52.94% of participants do not inform themselves about science regularly, even if part of them, curiously, work in science. However, 64.70% of them have contact with nature regularly. In general, their perception of science is very high with respect to the contribution of science to the well-being of society, the impact of learning science to understand the world we live in, and the evaluation of the effects of science in society.

We study the participants who actively took part in the use of the Natural Patterns (n=8). As far as science involvement is concerned, no significant differences are observed between before and after participating in Natural Patterns. The only noteworthy fact is that participants who had never participated in scientific activities expressed willingness to do so in the future. Regarding science perception, neither were there significant differences between before and after the use of Natural Patterns, just slight differences. In general terms, we could say that science disposition has not changed significantly due to the use of Natural Patterns. In fact, the participants had a high science disposition from the beginning.

2.5.4 General experience in the use of Natural Patterns

After the participants took part in Natural Patterns (n=8), they answered some questions about the whole experience. Some participants described the experience in terms such as: interesting (37.5%) or fun (20%): *"It was interesting to see the relation of patterns in the nature"* or *"I really enjoy with this experience because it makes me look around and see nature in other way"*. The most repeated reaction pointed out that the experience incentives a focus on natural surroundings (50%): *"It was really fun. It made me take my time and look to my surroundings and observe the wonders that nature creates every day and we don't pay attention to."*. Nevertheless, some participants reported problems with software or lack of clarity in the game.

Concretely, the vast majority of participants (87.5%) reveal that they enjoyed the experience of Natural Patterns. The same proportion of participants would participate again in this study and they would use Natural Patterns without financial reward but with some conditions: *"I will use it but not as much as I did during the study."* or *"Yes, but not every day."*. Along the same line, most participants (62.5%) considered that they had learnt about science collecting samples and performing the micro-task implemented.

A high proportion of participants (75%) consider Natural Patterns a game. Two participants do not clearly consider Natural Patterns a game, they point out: *"It's sort of a game – because*

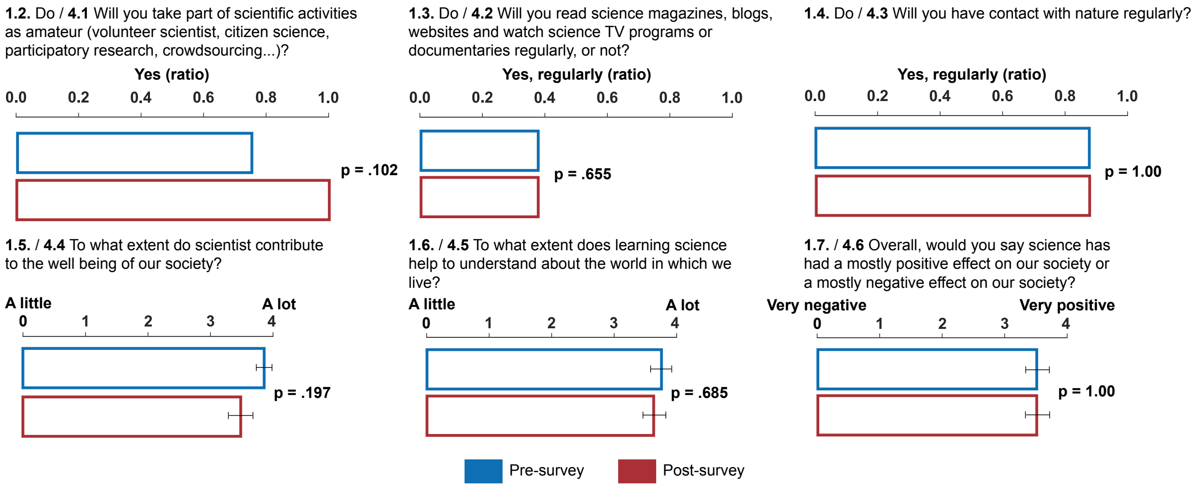

Figure 2.7: Science's involvement before and after Natural Patterns. Ratio of people (n=8) that answered affirmatively to the questions related with involvement in science before and after using Natural Patterns. χ^2 differences test before and after taking part in the experience. **Science's perception before and after Natural Patterns.** Science perception of the participants (n=8) before and after the use of Natural Patterns. χ^2 differences test before and after taking part in the experience.

as you complete challenges and capture patterns you win points, so, in a certain way, it can be considered a game." and *"Not really – I've found the app could be a little more guided."*.

2.6 Discussion

The Natural Patterns case study sheds light on the needs that arise when we design crowdsourcing activities in citizen science. Here, we discuss the results in relation to the design implications that promote and encourage science disposition, the understanding of the scientific method, critical thinking, and finally, we explore the motivations. Furthermore, we conclude with some notes about the scientific method and science disposition implications.

2.6.1 Design implication

Natural Patterns picks up the design principles of a platform addressed to objectives that go beyond the scientific outcome. The core objective from the point of view of the user is to collect patterns of nature, and for this purpose we propose activities that not only promote the collection of samples, but also other actions.

Observing and discovering is crucial to capture samples. In *Captures*, from where a majority of contributions are made, we obtain a high interaction with nature and a careful observation of samples thanks to the action of taking and classifying patterns' photographies. Therefore, in this sense it meets the design objectives. However, the participants demand more details about the patterns, to expand the information and receive feedback, either from their partners or from the application itself. This point can be extrapolated to other collection projects where the participants are motivated to learn and therefore should be given information about the actions which they are taking, should have clear goals and their contribution should be valued beyond the simple collection action.

Beyond collecting samples that the participants find in a casual way, in *Challeges* the collection of specific samples is encouraged, for instance the collection of clouds. In this way, besides being able to go into detail in some very common patterns in nature, we can promote learning through repeated observation. However, the evaluation in terms of learning was not positive and, although some challenges were very active, the participants wanted more interaction with the other participants by, for instance, adding scenarios of discussion and assessment to encourage socialization and competition respectively.

With *Hypothesis*, we go further by building an activity that allows participants to create hypothesis, their own micro-taks. In this case, the participation was positive and the creation of hypotheses encouraged the creation of more hypothesis. Thanks to the cooperation of the other participants, their contributions, the discovery of new patterns was promoted. Unlike other sections, socialization and discussion go hand in hand in *Hypothesis*. Nonetheless, some participants asked for more interaction among them to discuss the hypotheses and their contributions.

Interaction with nature

One of the fundamental objectives of Natural Patterns is to make the participants interact more with the environment that surrounds them and specifically with nature. Natural Patterns achieves this goal, as most participants reported that their interaction with nature and the observation of their surroundings increased. Above all, the microtasks of *Captures* and *Hypothesis* have a high degree of interaction with nature and the surroundings.

Despite other microtasks such as *Challenges* and *Nearest Patterns* being designed with the aim of looking at the patterns that surround us, they have not had the expected impact. The comments received about *Challenges* noted that the objective is not very clear and the task's design is not engaging. Therefore, it is necessary to rethink how this micro-task is presented to achieve the objective we are looking for. *Nearest Patterns*, which is the worst rated feature, has a different interpretation which makes it clear that it is necessary to propose an alternative examination for this feature.

From the interaction with nature, a minimum knowledge is expected to emerge in order to understand how patterns are formed and under what circumstances. However, Natural Patterns seems not to achieve very high levels of understanding of nature. To accomplish that, we could design ad-hoc microtasks, for example in *Challenges* or *Hypothesis*. In fact, *Hypothesis* is the section in which the levels of understanding are higher.

Motivations

The general assessment of Natural Patterns is positive as regards the motivations of the participants and the perception of the game as another citizen science gamified projects (Ponti et al., 2018). Definitely, addressing a scientific activity in the terms of a game makes the perception of the activity motivating. In this version of Natural Patterns, the gamification layer is very basic, and it focuses on a scoring system and rankings, so there is further scope to improve the game system.

If we focus on the analysis of the motivations' indicators, we observe that they have acceptable scores in the main design features, but relatively lower in what refers to interaction with nature and the scientific method. *Hypothesis* is the section that has motivated the participants the most in all angles: engagement, socialization and, together with *Captures*, competition and reward. However, there is room for maneuver to improve the motivational factors that participants perceive.

Citizen science can learn from platforms that, in principle, are not designed with a scientific objective and that, by chance and thanks to their widespread use, are useful to achieve secondary objectives, for example in the conservation movement (Dorward et al., 2017; Khelifa, 2016; de Oliveira Roque, 2016; Nature, 2016b)

2.6.2 Scientific method

The scientific method and the process of scientific research can be defined, as stated above, in many ways. Nevertheless, all of them have a set of concepts in common: observation, analysis, collection, hypothesis or discussion, to mention just a few. Most participants in the study had basic notions of what the scientific method is. After the use of Natural Patterns, they valued the impact that the platform has on the understanding of the scientific method positively. From the results, it can be concluded that Natural Patterns helps in the observation, collection, analysis by comparison and in drawing conclusions. However, the communication of results and conclusions is not viable with the microtasks and design features implemented. Conclusions regarding the study of the features related to the scientific method are similar. The main features of the platform promote interaction with nature and observation, critical thinking and discussion. However, there is a lack of transmission of knowledge and socialization. We can introduce new features in the application that rectify this situation, adding microtasks based on the exchange of experiences. Ultimately, Natural Patterns creates an environment that largely represents scientific thinking, in tune with (Trumbull et al., 2000).

2.6.3 Science disposition

We study science disposition from two perspectives. One relates to the implication that the public has in scientific activities, passively or actively, and which we describe under the idea of science involvement. The other relates to the perception that the public has about science, which we encompass under the term science perception. The results of the study show that, in general, science disposition among the participants is high, especially in relation to science perception rather than in science involvement,which shows slightly lower ratios. After participating in the experience of Natural Patterns, no significant changes in their perception are observed.

On the one hand it should be noted that, as in many other citizen science projects, the participants are people who already have a good disposition towards science (Price and Lee, 2013). The really complicated thing is to attract people who do not have that initial disposition. In future studies it would be positive to introduce new audiences who would not engage in scientific activities on their own initiative (Bonney et al., 2015). This could be accomplished through community partners, by developing activities in specific and diverse contexts (e.g. schools, institutes, civic centers, museums, etc.) or using online platforms for recruitment (e.g. SciStarter).

2.7 Future works and Limitations

As a result of the study and the natural evolution of crowdsourcing systems, a set of lines are drawn up to move forward and correct certain limitations. Here, we describe the limitations and possible lines of action.

Firstly, a long-term longitudinal study without incentives is necessary to deepen into some preliminary findings presented in this work, especially with regard to science disposition. The study of the science disposition serves as a reference to assesses the thoughts of the participants in the study, but it does not assess changes due to the participation in Natural Patterns. Especially because it is a short study which does not allow for the behavioral changes to emerge.

In fact, the sample has a high disposition toward science, and this can bias the observations about the motivations, positively or negatively, depending on their expectations. Consequently, it is necessary to continue researching science disposition and the creation of activities that promote a positive attitudinal change on the perception of science. It is also necessary to encourage participation in scientific activities that allow scientists to approach the general public and that the public feels part of the scientific activity.

We need also consider that the micro-task proposed by *Nearest Patterns* can not be rigorously studied since it is fundamental for the participants to be close to each other. This means that other types of studies have to be planned for this micro-task in which the participants have to be in the same physical context (Harburg et al., 2015). The incorporation of notifications, especially to alert of interactions with patterns that are close, can be an interesting alternative to further enhance the relationship with the environment.

As noted above, there is a need, related with nature understanding, to improve explanations about the influence of basic sciences in the emergence of patterns in nature. It is essential to explain the bases of many ordinary natural patterns, on the platform itself or using an alternative system (web, video channel, etc.). Create a pool of questions about patterns in which participants during a period time contribute with his answers could be a new micro-task that engages participants similarly to *Hypothesis*.

It would be advisable to introduce the co-design in Natural Patterns in a much deeper way (Senabre et al., 2018; Shirk et al., 2012; Delfanti, 2010; CAISE Inquiry Group, 2009). In particular, now co-design is intrinsically implemented in the *Hypothesis*, the high rated model, where participants create their own hypotheses (microtasks) and the other participants contribute to them. However, co-design can be a good method to create new microtasks but also to create activities at a higher level, always following the basic foundations of Natural Patterns: the natural patterns, the scientific method and the interaction with the environment, while addressing the interests of the participants or other collectives (NGOs, local communities, etc.).

In order to further promote engagement, the game layer could be expanded. Segmenting the activities and creating badges for the main micro-task, as well as, giving more functionality to players with more experience than others, is also a possibility. These could be some alternatives that increase the playability, and therefore the motivations of the participants.

In principle, Natural Patterns is intended for the creation of scientific activities that do not have a primary object in the investigation of the samples collected, but rather in the activity itself, as occurs in other projects (NASA, 2017a,b). However, the idea is to share the data in open repositories so that other scientific projects, specially on natural science, can benefit from it.

In Natural Patterns the samples are captured and classified by each participant. It is necessary to incorporate basic features such as the multi-tag classification, a posteriori information editing and evaluation of others' contributions. However, in order to achieve the whole experience of scientific thinking, it would be desirable to obtain all the knowledge about a natural phenomenon (in our case natural patterns) in the same way that it has been traditionally gathered in natural sciences: understanding, describing, quantifying, predicting and controlling the phenomena. One step further is to create an intelligent classification system, meaning that we train a system with samples of patterns to classify in categories. In addition, this system could be enriched with the new contributions of the participants (Dieleman et al., 2015; Rebecca Boyle, 2016; INaturalist, 2017; Emily Matchar, 2017).

Finally, there is the possibility of expanding the methods of data collection. In this case, it is possible to provide samples of sounds. Thus, we enrich the way in which the scientific method is applied, specifically in data collection, and at the same time we do not only focus on visual patterns, but we extend to sound patterns.

2.8 Conclusions

This work explores the design of a platform for citizen science crowdsourcing with the objective of captivating the public to interact with their natural environment, to explore, to discover, to be immersed in the scientific method and to boost critical thinking. The Natural Patterns platform conducts this interaction through observation, capture, collection, analysis and discussion of natural patterns that we can find in our, more or less natural, environment. We conduct a study to comprehend if the activities, microtasks and design features implemented help us to create the complete experience that we propose. Our findings suggest that the design of the platform promotes a high interaction with nature. The scientific method is mostly well represented in the absence of improvements in communication and interaction between participants. The motivations of the participants can be promoted by introducing more gamified design features or by co-designing activities or microtasks. Since the interest in the formation of natural patterns flowed naturally with the participation in the platform and, it would be interesting to introduce functionalities that allow a greater understanding of nature, as well as including new design features to enhance socialization and discussion. In general, the insights suggest that the platform is a good environment to promote the interaction with nature and enhance scientific thinking. Even so, we consider that the study has certain limitations, that it is necessary to carry out a longitudinal study to know the long term impact in science disposition, besides increasing the sample in size and diversity, especially introducing people with a low disposition towards science.

Chapter 3

Citizen Social Lab: A Platform for Human Behaviour Experimentation

Summary – Our approach to the performance of behavioural experimentation revolves around opening the experience to the general public, focusing on the participation and moving the experiments to the field, to the appropriate context where the behaviors naturally occur. We introduce a scalable platform with a suite of behavioural games and social dilemmas that helps researchers to perform behavioural experiments in the wild in order to study behavioural traits – cooperation, trust or collective sense –. It is designed to engage a more general population into behavioural experiments, thus broadening the results and obtaining more general theories about behavior by means of performing pop-up experiments.

3.1 Introduction

Social dilemmas modelled as behavioral games are important tools to study the general principles of human behaviour and to understand social interactions. Social dilemmas occur when individual interest conflict with other individual or collective interests. Behavioural experimentation thus yield relevant scientific outcomes that have been used to test theories and to refine models, providing experimental data for simulations (Sánchez, 2018), and making the understanding of human behaviour move forward. But the impact of the experimental insights go beyond the scientific theories, because social dilemmas describe interactions and conflicts in real-life situations such as climate change mitigation, refugee repatriation, utilization of public space, social inclusion, gender discrimination, care-in community in mental health or resource depletion, and results can be translated to improve all these areas.

Traditionally, most experiments have been conducted in laboratories with highly controlled experimental protocols but with limitations in terms of subject pool or decisions' context (Levitt and List, 2007b,a, 2008; List, 2009). There is a sample bias since most laboratory experiments' sample consists of students who have a socioeconomic and sociodemographic situation. Thus, those studies do not reflect the general population behavior (Levitt and List, 2007b; Fehr and List, 2004; Rosenthal and Rosnow, 1969; Orne, 1962). Besides, generalizability of results of laboratory experiments also is affected by the physical context in which they are performed. The situations of social interaction that are studied do not happen in laboratories, but in real life scenarios where participants face dilemmas and take decisions. This leads participants in laboratories to not engage in real-world behaviors, but instead in behaviors that are biased by the experimental conditions.

Furthermore, recently social experimentation has been affected by the general crisis of science in replicability and reproducibility, issues that concern the main actors in science (Chang and Li, 2015; Open Science Collaboration, 2015; Munafò et al., 2017). Some efforts have been done to solve this situation, promoting the transparency in the statistical and methodological aspects of laboratory work, but also promoting the publication of more detailed methods, the data sources and the codes used in the experiments and in the analysis (Nature Editorial, 2017a). Scientists are encouraged to conduct replication studies (Nature Editorial, 2017b) and, in general, to pursue a more open research culture (Open Science Collaboration, 2015; Nosek et al., 2015).

In recent years, Computational Social Science has emerged as a multidisciplinary field that studies complex social systems and provides new insights about social behaviour, combining tools and method from social and computer sciences (Lazer et al., 2009; Cioffi-Revilla, 2010; Conte et al., 2012; Mann, 2016). In this line, a large number of studies have been conducted generally exploiting big amounts of social data, mostly collected from online social platforms (Twitter, Facebook, Coursera, etc.). However, there is a missing scale gap between the studies conducted with large-scale data from online platforms (that comes from uncontrolled protocols) and the small-scale data collected from the experimentation in behavioural science labs (collected with robust protocols). New platforms fill this gap providing opportunities for the design of mid and large-scale behavioral experiments in online labs that guarantee the quality of the data collection (Radford et al., 2016; Chen et al., 2016; Holt, 2018). These more flexible platforms have great advantages, as they facilitate the recruitment of more diverse sociodemographics profiles or very specific communities according to the needs of the experiment, are able to carry out the experiments in a distributed way in space and time, and they are more efficient at the economic level, since the infrastructure is much lighter. However, other limitations arise such as the identification of the experimental subjects or the economic incentives, to mention only a few.

Our scenario of experimentation is described in the context of pop-up experiments (Sagarra et al., 2016), an intermediate situation between traditional behavioural experimentation and big data analysis. The basic idea is to translate the experiments outside the lab to real context, and to open participation to new and more diverse audiences. More importantly, the experiments are not only build taking into account the researcher's interests and motivations, but also considering the perspective of citizen participation and its social impact in terms of providing the right knowledge to conduct new evidence-based policies by public administrations and empower participants to trigger civic actions. This is framed within the citizen science approach (Bonney et al., 2014; Gura, 2013; Hand, 2010; Silvertown, 2009), that promotes the participation and inclusion of non-expert audience in real research processes in different ways (Kullenberg and Kasperowski, 2016; Bonney et al., 2015) (co-creating projects, collecting data, interpreting and analyzing data, and provide actions based on the evidences collectively gathered). Citizen science helps us to involve the general public in behavioural experimentation and impacts the participants themselves (Senabre et al., 2018; Bonney et al., 2015; Price and Lee, 2013; Bonney et al., 2009), for instance increasing their disposition to science (Perelló et al., 2018).

To carry out these experiments interactively, we designed and build Citizen Social Lab, a platform with a collection of decision-making and behavioral games, among which we can find a Prisoner's Dilemma, a Trust Game, a Dictator's Game or a Public Goods Game. Depending on the goal of the experiment and the behavioral variables to be studied, the researcher can select and parametrize one or various games, and also define the general dynamic of each game. The platform registers all the behavioural actions taken by the participants, but also provides surveys to collect sociodemographic data, information about the participants' experience or their decision making process.

Since the platform is mostly used in experiments performed real-life contexts, it only requires a light infrastructure that can be installed and executed in different locations in a simple and robust way. The platform does not allow the intervention of uncontrolled participants, and registers data accurately without alterations of any kind. In addition, unlike experiments in laboratories, in general the participants in these experiments are recruited at random. As they are not captive participants, both the experimental staging and the platform need features that engage people to participate. For this, different approaches are used, one of them being the gamification of the experience (Ponti et al., 2018; Bowser et al., 2013), which consists in presenting the experiment as a game and a scientific investigation at the same time. Another important ingredient and challenge is to increment sense of ownership of the data being gathered by for instance delivering back a personalized report to each participant and by organising public lectures to discuss aggregated results once the papers is been published. This effort add new dimensions to the always required open data access or ethical and transparency requirements when dealing with citizen science approaches.

In this article we present the experimentation platform that has been active since 2013 and has been used successfully and since then 15 experiments to study different aspects of human behavior. Up to this date 2821 people have contributed, taking around 45200 valid decisions.

3.2 Materials and methods

3.2.1 The platform

Citizen Social Lab is a platform designed to assist in the deployment of human behavioural experiments. It has been created with three important goals in mind that foster versatility. First, the platform is based on light and portable technologies, so it can be used in open and diverse environments following the guidelines of popup experiments (Sagarra et al., 2016), but also in pure online or in more “classical” experimental laboratories. Second, it has been designed with a friendly user interface to facilitate participation to a general population, and to engage and motivate participants to solve the tasks proposed in the experiment while they have an enjoyable experience. And third, it is structured in a way that it is easy to incorporate any type of social dilemma or behavioural game, as well as, any type of interaction: individual/computer, individual/individual and individual/collective.

The platform allows researchers to carry out a suite of dilemmas or behavioural games, which compose the core of the system. The system already contains a few different available dilemmas, which are described in the next section, and this number is expected to increase as new experiments are developed and deployed using the platform. Moreover, beyond the data collected from the participant’s decisions, the system is designed to collect complementary data (about sociodemographics, user experience or experiment-related questions) through surveys before or/and after the social experiment takes place. It also registers all the activity of the participants when using the platform, which can be used to infer other parameters (e.g. response time).

The platform architecture is highly modular and allows the researcher to construct personalized environments combining and parametrizing the modules they require for their particular experimental setting. The basic client modules currently available are the following: (i) Introductory interfaces, with brief but detailed information about the topic and goals of the experiment and legal information with privacy policy; (ii) Questionnaires; that can be used to collect sociodemographic information and also to present specific questions related with the experiment topic or setting. Questionaries can be used before and/or after the main experiment. (iii) Tutorial and instructions; so participants can learn the rules and the mechanics of the experiment by themselves (even though in the physical location there are always researchers to provide support if any question arises) and practice a few testing rounds of the game to familiarize themselves with the game interface. (iv) Games and/or dilemmas; the core of the platform, the module

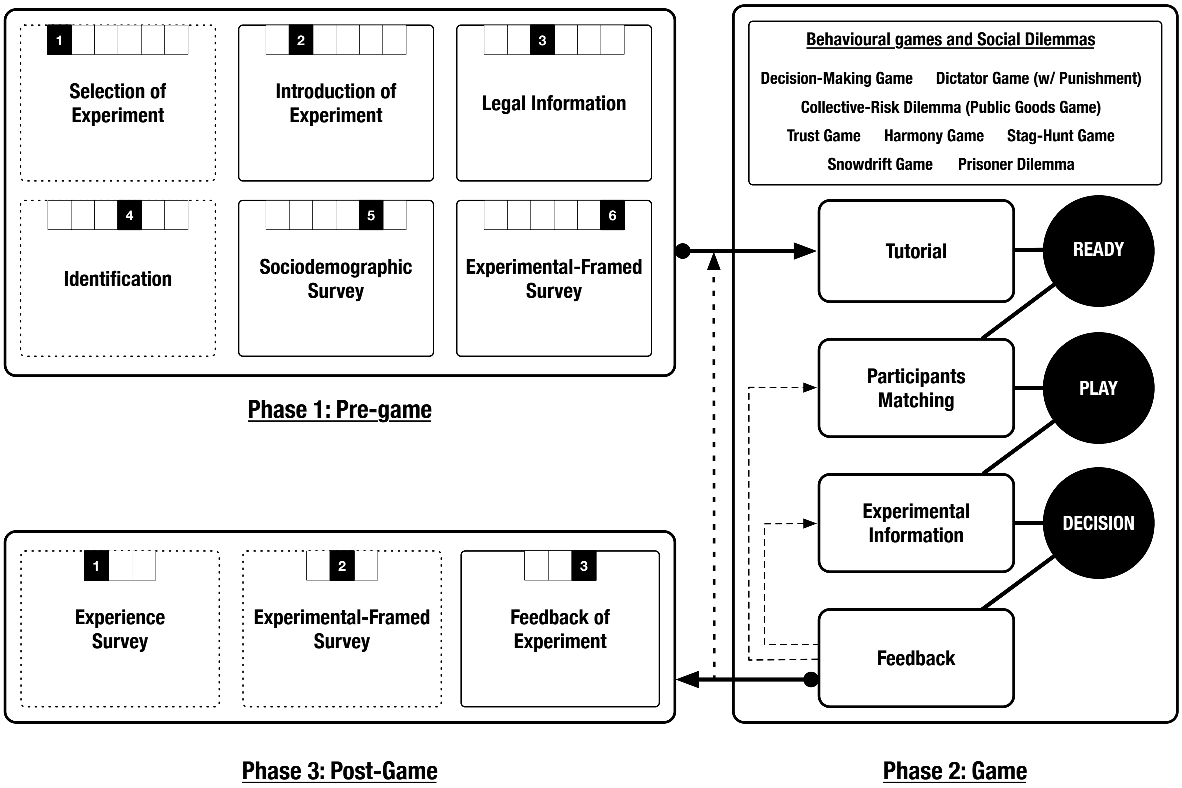

Figure 3.1: Block diagram of a participant's flow through one experimental setup. The participant goes through three stages: the first stage contains the pre-game module with preliminary instructions about the experiment and surveys, the second stage contains the core game mechanics (which implements the suite of decision-making and behavioural games), and the third stage consists of the post-game module with the final feedback of the experiment and surveys about the experience and the topic of the experiment. Not all these modules and interfaces are present in all the experimental setups.

that runs the experiment to collect the decisions of the participants. An experiment can incorporate only one game or a collection of them. (v) Results; a set of interfaces designed to provide feedback to the participants on the outcome of their decisions in the experiment. This is crucial to increase the positive return that the obtain for participating in the experience. Finally, (vi) the administration interface is composed by a set of pages that let the researcher to control the parameters of each session, monitor the evolution of a game, and overview the general performance during the experiment in real-time.

The modules are configured to define what we call the participant's flow through the experiment (see Fig.3.1). The system is designed to guide the participants through all the stages without the need of interacting with a researcher (unless otherwise required by the participant).

Games Module

The main goal of the platform is to collect the decisions of the participants when they face different types of dilemmas that are analogies of real-life situations. Most of the dilemmas included up to now are social, which require synchronized interaction with other individuals, however the platform can also be used to study individual decision-making situations that do not require real-time interaction with other participants.

The first social dilemma implemented is a generalized version of a simple dyadic game, where two people have to decide simultaneously which of the two actions they will select, and the outcome is the result of the combination of them. Depending on the values presented to the participants, the can face different types of games: a Prisoner's Dilemma (Axelrod and Hamilton, 1981a; Rapoport and Chammah, 1965), a Stag Hunt (JABONERO, 1953), a Hawk-Dove/Snowdrift (Sugden, 2004; Maynard Smith, 1982; Rapoport and Chammah, 1966) or a Harmony. These dilemmas can be used to measure two important features of social interaction, namely the temptation to free-ride and the risk associated with cooperation.

The second social dilemma, the trust game (TG), or otherwise called the investment game, is used in order to measure trust and reciprocity in social interactions (Berg et al., 1995). In TG two players are given a quantity of money. The first player send an amount of money to the second player, the first player is informed that the money that he sends will be multiplied by a factor (e.g. three). The second player makes the action of give some amount of the multiplied money back to the first player, and then both receive their final outcome.

The third social dilemma, the Dictator game (DG) can be used to measure rationally self-interest or distribution fairness (Fehr and Fischbacher, 2004). In this game, the first player "the dictator" splits an endowment between himself and the second player,"the recipient". Whatever amount the dictator offers to the second player is accepted, therefore the recipient is passive, cannot punish the dictator's decision. DG is not formally a game because the outcome only depends on the action of one player, in game theory those games are known as a degenerated game. However, there are a modified version of DG which includes a third player who observes the decision of the dictator and has the option to punish the dictator's choice. The third person receives an endowment that could choose to spend to punish the dictator, so that punishing has a cost for the punisher.

The fourth social dilemma, is a variant of the public goods game, which is a collective experiment game in which the players with their contributions decide invest in public goods or keep their private goods. This particular version is known as collective-risk dilemma (Tavoni et al., 2011; Milinski et al., 2008), and consist of a group of people had to reach a common goal by making contributions from an initial endowment. If the goal was reached, every individual received the part of the money not contributed. If not, a catastrophe occurred with certain probability, and all participants lost all the money they had kept.

The platform also includes a decision-making game, where participants have to take decisions having uncertain and/or incomplete information (Gutiérrez-Roig et al., 2016b). This game is played individually so there are no interactions with other players during the game. With this game we can study decision making strategies by controlling the type and amount of information that can be accessed by the participants.

All the dilemmas described previously can be parametrized to allow for different types of studies (for instance, controlling the values of the payoff matrix) or extended to include different variations when they are available. Also, starting from the implemented interaction structures (Fig.4.2), new dilemmas can also be constructed and added to the platform following a simple set of guidelines described within the code of the platform.

Participation and Motivations

Moving the experiments out of the laboratories implies that the participants are not captive in advance. The platform is designed to be used in open environments, and attract new audiences. The recruiting process in open environments such as a games' festival or public space is substantially different from the recruitment in laboratories, and presents more challenges. In the pop-up experimental framework (Sagarra et al., 2016) we usually include a narrative context and performative elements to attract the attention of the participants. However, it is also important once the attention of potential participants has been attracted to present the experiment in a motivating way.

We use gamification techniques insofar the experimental settings allow us to ensure the scientific rigour of the experiments. Behavioral games and dilemmas per se already have elements and mechanisms of games such as: challenges, objectives, rules, reward, punishment, interaction, competition, collaboration, call-to-action, among others. Based on them, we create an experience where we present some of the experiments as games, with a narrative setting that creates a story surrounding the experiment. In some experiments, mainly the ones that took place within the DAU Festival, an actor is in charge of the recruitment characterized as the main character of the experiment.

The experiments are designed to enhance the motivations of the participants, not only from the perspective of games, but also to impact in the science disposition of participants, the understanding of science or the impact in social issues. This is the case of the framed experiments: The Climate Game, Games for Mental Health, the games for social change within the STEM4Youth project and the street art performance called urGentEstimar, all of them are focused on real social concerns: a collective climate action, the mental heath promotion of in-community care services or the concerns fom several school groups related to social inclusion, use of public space and gender violence. Furthermore thanks to the participants' contributions, they participate in the advances of science, they receive, an economic incentive to participate (according to their performance in the game).

In our case, there are two types of participation according to experiments' context. Most of them have been carried out in uncontrolled environments in terms of recruitment, without "captive participants" (e.g. festivals or public spaces). On the other hand, in specific cases, where the experiments were carried out with local communities, they have been developed within the community, therefore, in this case, the need to apply special recruitment techniques is not so important. In any case, to support the game-based approach, the platform allows the introduction of resources to include the narrative, always preserving the scientific rigour, and also provides features that can be used to create a gamified experience.

Technical Details

Some of the dilemmas previously explained require of individuals interacting in different manners. For instance, in games where two individual participate there are at least two possible interaction styles: one where the two individuals make a simultaneous decision without knowing what the other decides and after that they receive the outcome; or another where one player makes a decision while the other player is waiting, once the first decides the second, knowing what the other decides, makes his decision, finally both get their final feedback (see Fig. 4.2). Also, experiments can have different evolution mechanics: from one-shot games, in which the players just make a unique decision, to iterated games in which the players make various decisions consecutively with the same or different participants. And finally, we also have to consider the possibility that the interaction between the players can be constrained by an underlying structure that defines the relationship between the players, which can range from a all-connected-to-all structure to a specific network structure.

Taking all these points into consideration, we designed a client-server architecture that controls the flow of the experiment. On one hand, the server manages the pace of the experiment, and implements all the core games and synchronization methods between players. It is based on a python-django backend, combined with a database to store the information generated separately by each experimental setup. The server can be run online, to allow experimentation on the internet or it can be installed in a local server to run experiments in local area networks.

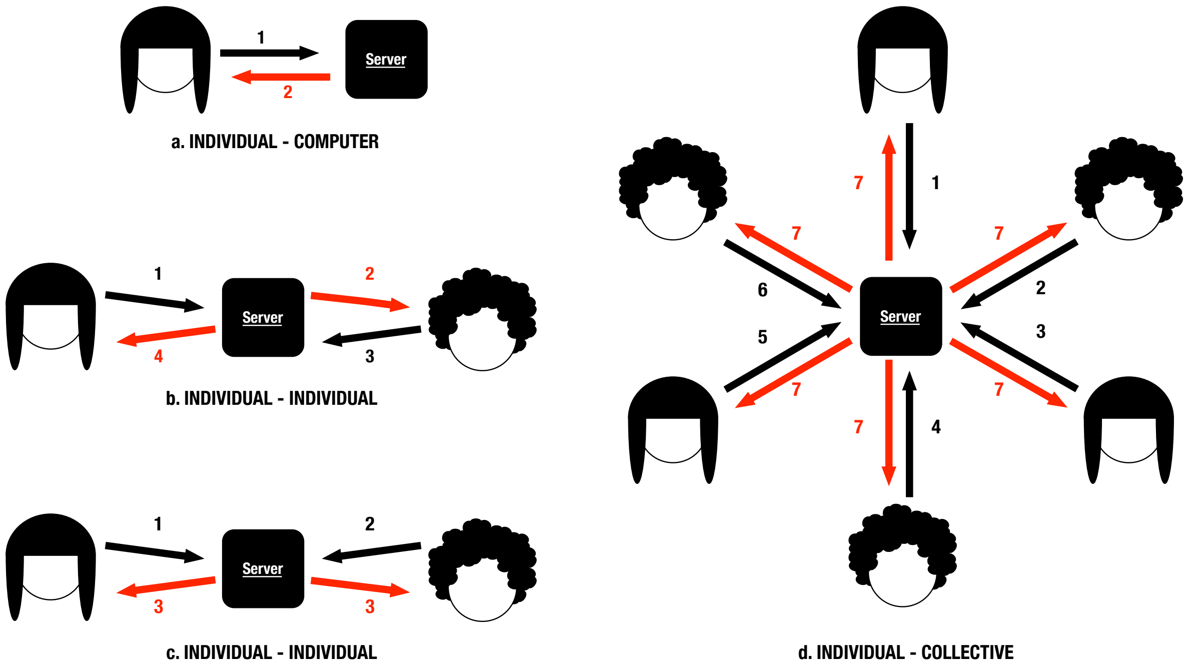

Figure 3.2: Interaction types included in the platform. The platform currently implements four different types of interaction that cover individual-computer (a), individual-individual (b, c) and individual-collective (d) types of coordination.

On the other hand, the client contains the user interface that the participants have to use to interact with the experiment. The technology on the client side is composed of html and javascript files that are generated dynamically from the experiment description files. The user interface has been designed to fit the resolution of a tablet device, but also works with any computer with a standard browser.

Most of the experiments have used the same infrastructure consisting of a laptop that acted as a server and a collection of tablets that allowed up to 30 participants to be simultaneously participating in the experiment. In Fig. 3.3 we present a diagram of this infrastructure. Data is collected and stored in a database (which may be relational or not), and personal information is stored separately from the experimental data to follow the privacy guidelines required by this type of experiments.

Finally, live control of an experiment is critical to guarantee its correct development. For this reason, they can be controlled using an administration webpage that provides two features: it allows the researcher to configure the parameters that will be used in each iteration of the experiment (e.g. select if a certain group will be intervention or control) and it presents interfaces with the status of the experiment. Live monitoring can be done at two different scales, at a particular game level, where researchers have real-time detailed information about the evolution of a particular game (rounds played, decisions made, earnings, connection status, ...), or from a more general point of view to obtain a summary of the status of the experiment (demographics, games played, global earnings, ...).

Also, another important minor feature that we have taken into account is the internationalization of the interface, so each experiment can be easily translated to other languages.

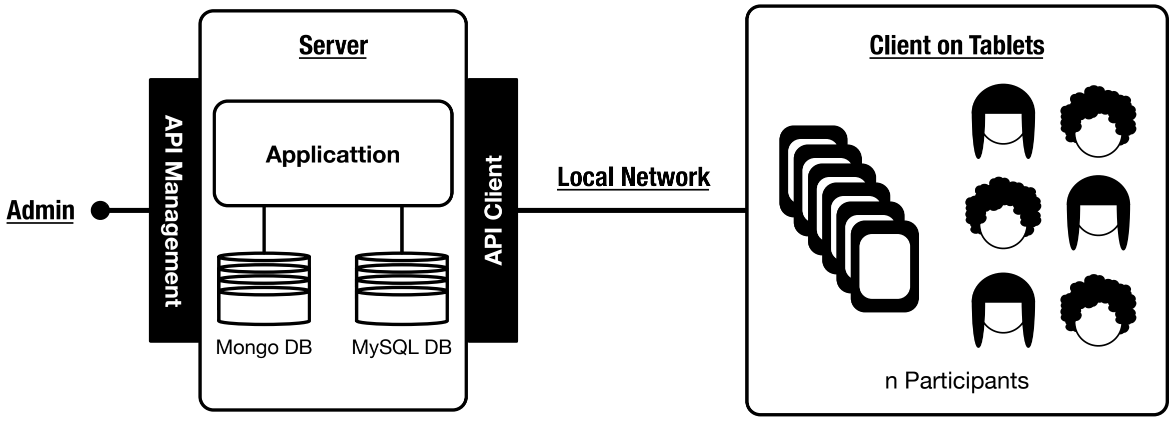

Figure 3.3: Example of the platform infrastructure. This is the basic technological infrastructure used in the majority of experiments. It i designed to be rapidly deployed in any environment.

3.2.2 The experiments

The platform has been in use since December 2013 in 6 different experimental setups focused on the analysis of human behavior. Some of them have been repeated in different situations, which adds to a total of 15 experiments realized. In this section we describe the main goals and results of the six research projects based on this platform, which are also summarized in Table 3.1.

1. The first experimental setup based on the platform is “Mr.Banks: The Stock Market Game” to study how people make decisions when they have limited and incomplete information. This setup emulated a stock market environment in which people had to decide whether the market would rise or fall. It allowed us to study the emerging strategies and the relevant use of information when making decisions under uncertainty, and the results are published in (Gutiérrez-Roig et al., 2016b). Three experiments based on this setup have been done in different locations, and is now available online [1].

2. Next, we created another experimental setup entitled “Dr.Brain" to study the existence of cooperation phenotypes. The games played by the participants were based on a broad set of dyadic games and allowed us to deepen our understanding of human cooperation and to discover five different types of actors according to their behaviours (Poncela-Casasnovas et al., 2016).

3. The following experimental setup included in the platform was “Dr.Brain: The Climate Game”, which was based on a collective-risk dilemma experiment to study the effect of unequality when participants face a common challenge (Vicens et al., 2017). Results showed that even though the collective goal was always achieved regardless of the heterogeneity of the initial capital distribution, the effort distribution was highly inequitable. Specifically, participants with fewer resources contributed significantly more (in relative terms) to the public goods than the richer - sometimes twice as much.

4. The fourth experimental setup implemented in the platform was called “Games for Metal Health" which was repeated in 4 different locations. The goal of this project was to

1. http://www.mr-banks.com

evaluate the importance of communities for effective mental health care by studying different behavioral traits of the different roles of the ecosystem. The results presented in (Cigarini et al., 2018b) reinforce the idea of community social capital, with caregivers and professionals playing a leading role.

5. In the context of the EU project STEM4Youth we performed three experiments, which were co-designed with high-schools of Barcelona, Badalona and Viladecans. They addressed topics raised in workshops with students: gender inequalities, use of public space and integration of immigrants. The experiments combined a set of games that included Trust Game, Dictator's Game, Prisoner's Dilemma and Public Goods games.

6. Finally, we performed two experiments named "urGentEstimar" in the context of artistic performances in Tàrrega and Poblenou (a Barcelona neighborhood), in which the participants took part in a set of behavioural games which included Prisoner's dilemma, Dictator's game or Snowdrift, and which were framed around different concerns of local communities.

3.3 Platform evaluation

In this section we analyze the versatility and the robustness of the platform by reviewing some of the results obtained by its use in different experimental setups. Mainly we focus on the sociodemographic diversity, the experience of participation, the time response data collected in the iterative experiments, and finally the robustness in the replicability of experiments.

3.3.1 Sociodemogaphic

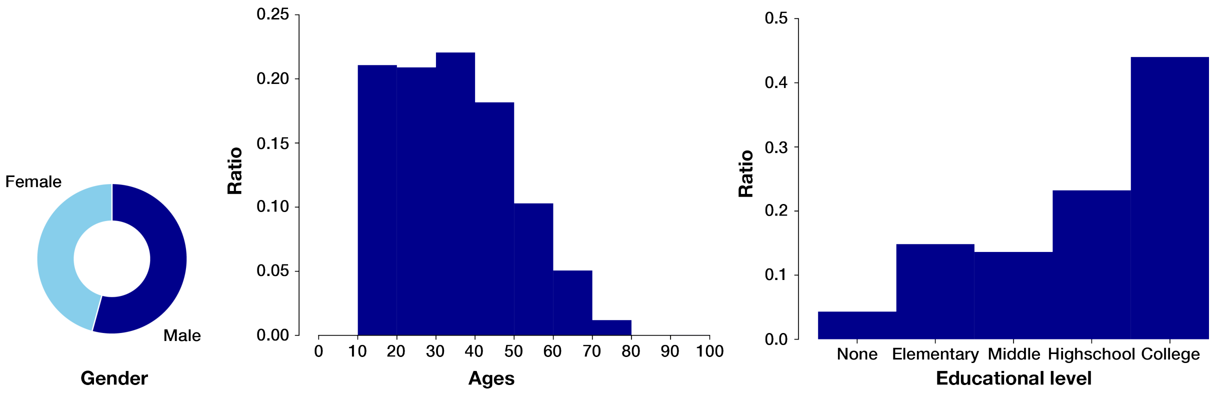

Figure 3.4: Diversity of the participants pool. (Left) The proportion of participants in all the experiments (n=2821) regarding gender is 54.27% males and 45.73% females. (Center) Distribution of participants according to their ages in all the experiments (n=2821). (Right) Educational level of participants in all the experiments except "urGentEstimar", which didn't ask this question to participants (n=1993).

To start, we review some of the demographical data of the participants in the different experiments. We already stated that one of the main goals of the platform was to open the experiments to a more general population. In this direction, in Fig.3.4 we present an overview of the 2821 people that took part at some point in the behavioural experiments and perform the experiment with this plarform. We observe that we had a combination of participants from a wide range

Table 3.1: Summary of experiments performed thus far. The suit of games is formed by: Decision-Making Game (DM), Harmony Game (HG), Snowdrift Game (SG), Stag-Hunt Game (SH), Prisoner's Dilemma (PD), Trust Game (TG), Dictator's Game (DG) and Collective-Risk Dilemma (CRD). The number of participants and decisions are the valid ones.

Experiment	Location	Date	Games	Participants	Decisions	Publication	Data
Mr.Banks: The Stock Market Game	Barcelona	Dec.2013	DM	283	18525	Gutiérrez-Roig et al. (2016b)	Gutiérrez-Roig et al. (2016a)
	Brussels	Jul.2015		37	2397		
	Barcelona	Jun.2015		20	1078		
Dr.Brain	Barcelona	Dec.2014	HG, SG, SH and PD	524	8366	Poncela-Casasnovas et al. (2016)	Poncela-Casasnovas et al. (2017)
Dr.Brain The Climate Game	Barcelona	Dec.2015	CRD	320	3200	Vicens et al. (2017)	(Embargoed)[1]
	Barcelona	Dec.2015		100	1000		
Games for Mental Health	Lleida	Oct.2016	CRD, TG and PD	120	1680	Cigarini et al. (2018b)	Cigarini et al. (2018a)
	Girona	Mar.2017		60	840		
	Sabadell	Mar.2017		48	672		
	Valls	Mar.2017		42	588		
STEM4Youth	Badalona	Apr.2017	DG, TG and PD	151	1510	(In preparation)	(Embargoed)[1]
	Barcelona	Sep.2017		126	1260		
	Viladecans	May.2017	CRD	162	1620		
urGentEstimar	Tàrrega	Sep.2017	DG, SG and PD	756	2314	(In preparation)	(Embargoed)[1]
	Barcelona	Oct.2017		72	136		

[1] Embargoed until scientific publication.

of ages, specially from 10 to 50 but the greater too, and diverse educational levels, with a predominance of those with higher education. Gender is also balanced (45.73% females) compared with other similar experiments which are usually performed by students with sociodemographic bias.

3.3.2 Response times

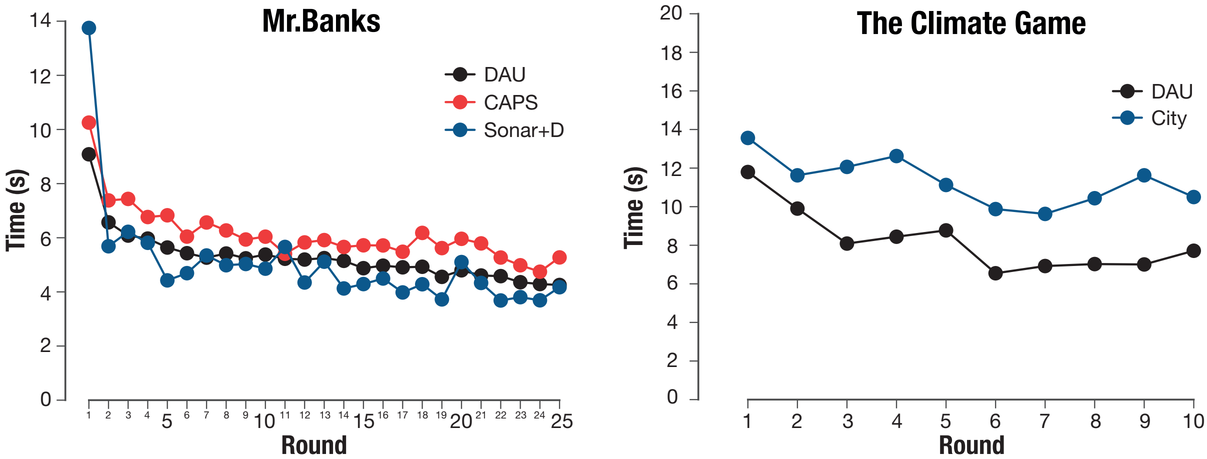

Figure 3.5: Time of response in different games. (Left) Time response evolution across rounds in Mr.Banks experiments for the main performance in DAU (n=283) and the two replicas CAPS (n=37) and Sonar+D (n=20). (Right) Time response evolution across rounds in The Climate Game experiment in both performances, DAU (n=320) and City (n=100)

The platform allows for the collection of very precise parameters about the participation in the experiments. One of them is the timestamp in which the participants perform an action. In iterated experiments, where participants make several decisions consecutively, the decision times are collected in each round so that we can calculate how long each participant takes to make a decision. An interesting parameter in behavioral experimentation is the learning time, or in other words, the evolution of time across the game.

In Fig.3.5 we can see the evolution of the decision-making time across rounds. On the one hand, Mr.Banks presents the evolution of the three experiments that were carried out, the main one (DAU) and the two replicas (CAPS and Sonar+D). The evolution of the time response during the three experiments shows very similar trends. In the first round the time is substantially higher than the rest of the rounds and we see that from the 5th round the slope softens and stays more or less constant until the end. In this experiment, the variables that come into play to make a decision are the same round after round, so the trend is maintained during the game. The three experiments show similar trends but slightly different absolute values; the context, size and heterogeneity of the sample may be the cause of this variation, which confirms the accuracy of the data collected.

On the other hand, in the case of The Climate Game the evolution of the game is somewhat different. The game starts with long times that go down gradually; however, depending on the point of the game in which the participants are (i.e. the distance to the goal) the times increase or decrease. In this case, unlike the previous one, the decision at each moment is given by the distance to the final goal, so that, as they approach to the end of the game, the times increase again. Therefore, here we observe two sets of behavior: the learning at the beginning of the game and the uncertainty as the participants reach the last rounds. The trends of the two climate change experiments are similar, however, the absolute value of time is slightly higher in the "City" context.

3.3.3 Robustness of replicability

We also measure the consistency and the robustness of the results across different repetitions of the same experiment. Some of the six experimental settings described in the previous section were repeated in different environments and locations, in some cases with similar populations (e.g. the mental health experiment) and in other cases with different populations (e.g. the mr.banks experiment). We focus on Mental Health and Mr.Banks to examine the robustness on the platforms in order to collect quality data allowing the replicability in different situations.

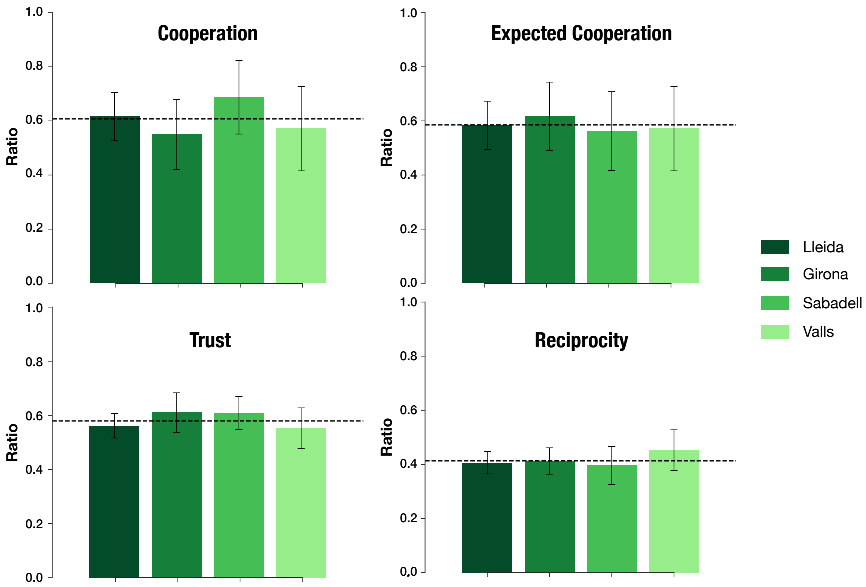

Figure 3.6: Robustness of generalization in Mental Health experiments. Levels of cooperation, cooperation expectation, trust and reciprocity in the four experiments: Lleida (n=120), Girona (n=60), Sabadell (n=48) and Valls (n=42). It is represented the average level with 0.95 CI in each case. The dashed line represents the total average levels. There are no significant variation in the level of cooperation (Kruskal-Wallis, H= 2.38, p = 0.50), cooperation expectations (Kruskal-Wallis, H= 0.38, p = 0.94), trust (Kruskal-Wallis, H= 2.67, p= 0.45) and reciprocity (H= 3.02, p= 0.39). See Ref. (Cigarini et al., 2018b) for further details.

Mental Health's experiments took place in Catalunya, in four different locations and social events (popular lunch, snack, etc.), in sum participated around 270 people. We analize the differences between the four events in cooperation, expected cooperation (Prisoner's Dilemma) and, trust and reciprocity (Trust game). The differences among the experiments in the four locations are not significative and the data can be aggregated to be analized as a whole 3.6.

Mr. Banks' experiment was performed in a main location, the DAU Festival, with a large participation, 306 people (283 valid participants), and obtaining robust results. From the analysis of decision emerged two strategies Market-Imitation and Stay-Win Switch-Lose. We compare the main result with two replicas that took place in two different events in Brussels (CAPS conference) and Barcelona (Sonar+D) with data from a narrow demographic populations and with the number of samples much lower than the main experiment. There are no significant

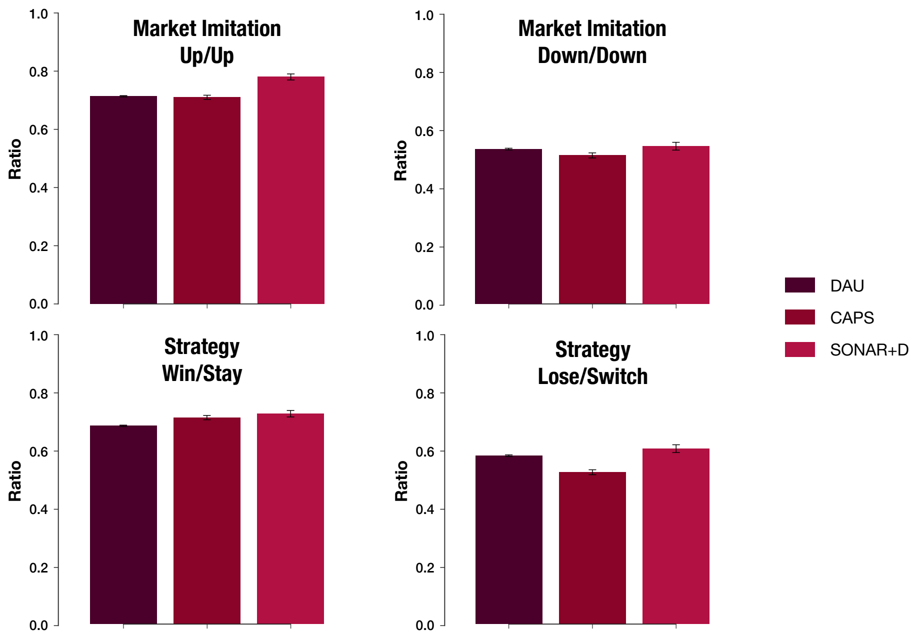

Figure 3.7: Stability of strategies in Mr.Banks replication experiments. Ratio to follow strategies of Market Imitation and Win-Stay Lose-Shift in the experiments: DAU (n=283), CAPS (n=37) and Sonar+D (n=20). There are no significant differences in Market Imitation strategies except the probability to Up/Up between the experiments of DAU and Sonar+D in (-2.53 SD). There are no significant differences in Win-Stay except in the last case (Lose-Switch) between the experiments of DAU and CAPS (2.35 SD).

differences (>1.96 SD) between the main experiment and the replicas except in Market-Imitation Up/Up between DAU and SONAR+D and Lose-Switch strategy between DAU and CAPS as Fig.3.7 shows. This mean that the platform capture data accurately, we are able to observe that the behavioural patterns found are consistent with the main results, because part of the results arise significant differences and the rest do not depend on the conditions.

3.3.4 Experience

Finally, another important aspect that we measured is the overall satisfaction of the participants after they finish the experiment. In the post-game survey of three games (Mr.Banks, Dr.Brain and The Climate Game) we asked the participants their level of satisfaction of the overall experience. Results of this question are presented in Fig.3.8. In all the experiments participants were mostly very satisfied or satisfied after the experience, specifically 82.77%.

3.4 Discussion

With Citizen Social Lab we present a platform that combines human behavioral experiments with a citizen science approach to attract a more heterogeneous audience and introduce science in the public context beyond the laboratories. The platform is designed to be signing a versatile, easy-to-use and robust, and it is applied in open and diverse environments. It has already been adopted in several experiments by thousands of participants from a wide range of demographics, which mostly valued to experience to be very positive. The results obtained by some of the

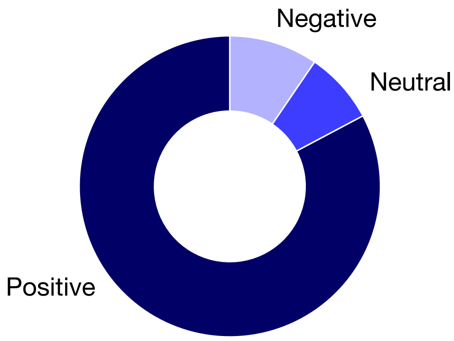

Figure 3.8: Participants experience. Experience of participation in Mr.Banks, Dr.Brain and The Climate Change (n=1178). The most of participants (82.77%) had a positive experience and a small group (9.51%) had a negative experience, the rest (7.72%) has an indifferent experience.

15 experiments realized with the platform have also shown the scientific validity of the data obtained from the platform with several scientific contributions.

In order to maximize participation and make it much more diverse than usual social experiments, we move the laboratory to the wild. In this non-friendly context we use the the pop-up experimental setup to draw the attention of the potential participants (which are all the people of the surroundings) with different techniques described in Sagarra et al. (2016). Then, we benefit from the lure of the game-base mechanics included in the platform in order to introduce them in the experience and guide them through all the tasks required by the experimental setup. This approach has proved to be very particularly successful in environments where they are likely to play (as the case of a games festival) and where the game-based experiments registered a high participation.

In this line, it is important to emphasize the need to adapt the experiments to the environment where they take places (e.g. musical festivals, scientific conferences, and so on), especially the experimental design and the interaction with the platform, because it is the way to make the empathy with potential participants increase. To achieve this, all the mechanisms of behavioral games and social dilemmas can be used to convert the interaction with the platform into a game (or other mechanism that fits in the context), always with the constraints imposed by the experimental scientific rigor. Based on the resulting scientific publications (Gutiérrez-Roig et al., 2016b; Poncela-Casasnovas et al., 2016; Vicens et al., 2017; Cigarini et al., 2018b), it is possible state that such an open and flexible platform allows to obtain robust and general knowledge on social interactions and human traits.

After all the experiments and their repetitions we consider that the platform has already reached a high maturity level, but there are several points that still need some work to keep improving the technical and experimental parts. First, the platform has been largely tested within the pop-up experimental setting in physical environments. However, even it has been designed to be easily integrated with online recruiting systems (e.g. amazon turk), it has not properly tested and validated in this environments. There is an opportunity to repeat some of the experiments to extend the consistency of the results when the dilemmas are presented to a purely online community.

Moreover, the platform is also constantly improving to provide new features and social dilemmas for the researchers. For example, we are creating the capacity for participants to create a unique profile and join in different environments. The long-term goal is to create a community of volunteers that participate in the experiments, and that can receive alerts when new opportunities to participate are open. We are also extending the number of available dilemmas within the platform as new research projects emerge which, once programmed and tested, are included in the main collection of available dilemmas.

The conceptual design in both types of experiments, the pop-up ones that have been done so far and the large-scale ones that are planned to be can be carried out in the future, have in common that the motivations of participants and scientific rigor are at the center of the participatory design. The platform has room for improvement in motivating the participants and in offering rewards at the level of learning and participation. On one hand, it is necessary to improve the mechanisms of learning about the scientific topic of experimentation during the participation in the experiment, but also about the nature of their contributions and about the positive impact in carrying scientific knowledge forward. In this sense, many experiments are framed within a context of social impact, so participation can also be associated to a call to action to solve social concerns. In the most recent experiments, this type of actions have been carried out outside the context of the platform, however, the online version can also contribute to this mission.

On the other hand, participants can improve their experience at the end of the experiment, not only receiving the necessary economic incentive but also obtaining an on-site feedback expanded with real-time information about the research process in which they have participated. They can also improve the experience by remotely following the evolution of the scientific research and participating in more phases of the scientific process. Another possible avenue to improve the platform is to build effective and real-time tools attached to experiments. Participants could in this way provide more feedback and actively contribute in the data interpretation and knowledge building process in both individual and aggregated levels. This effort appears to be meaningful to increase the participants' sense of ownership of the knowledge being produced by means of citizen science strategies.

With this article we also want to release the project to the researcher community so they can use it to create their experiments using the templates and guidelines already established in the platform. The project code is going to be released under a CC BY-NC-SA license. In the very end, if we aim to practice citizen science, it is also necessary to claim for opening the platform by all means: releasing data and code and opening up the results to make them understandable for anyone.

Part III

Human Behaviour Experimentation

Chapter 4

Mr.Banks and Dr.Brain: The First Experimental Performances

SUMMARY – Mr.Banks and Dr.Brain were the first two experiments that we carried out with the behavioural experimentation platform. In the first, participants had to predict stock market movements (up/down) under different circumstances; while in the second, four social dilemmas with different degrees of tension for cooperation were raised (cooperate/defect). Here, we will briefly present both experiments, the characteristics of the platform in these particular contexts, the main results, the analysis techniques and the robustness of the performance that lays the foundations for our experimental approach.

4.1 Introduction

In many everyday-life situations we have to take decisions, individually and collectively, with different degrees of uncertainty and information. In general, these actions are carried out in social contexts, so the decisions we make can also affect others. Therefore, there are interactions that make the outcomes depend not only on the individual decisions, but also on the collective ones.

The first experiment that we carried out using the platform for behavioural experimentation (described in Chapter 3) simulates the decision-making process in stock markets (Gutiérrez-Roig et al., 2016b). This context is interesting to understand how people make decisions in a scenario of uncertainty (Bell et al., 1988; Tversky and Kahneman, 1974; Pratt et al., 1964). There are many studies on trading strategies, price dynamics or uncertainty in the markets (Preis et al., 2013; Alanyali et al., 2013; Farmer and Lo, 1999; Samuelson, 1965). This, the stock market, is a framework that provides many ingredients to situate as people make decisions with different levels of information, and therefore uncertainty.

If we focus on the study of social interaction, meaning how people behaves when interacting, game theory and social dilemmas provide tools to study the tensions that exist between collective interests and individual interests, or conflicts between the rational and the irrational (Van Lange et al., 2013; Kollock, 1998), or the coordination and anti-coordination between the parties (JABONERO, 1953). Using the game framework we can approach the decision making process in a simplified way and study the cooperation.

To understand how we behave in situations of decision-making, we propose dyadic games – prisoner's dilemma, stag-hunt, snowdrift and harmony game (two-person and two-action) in which the participants have to decide simultaneously between two actions: cooperate or defect. In this way, we can study important characteristics of social interaction, like the temptation to free-ride or the risk of cooperation (Camerer, 2003; Rapoport and Guyer, 1966).

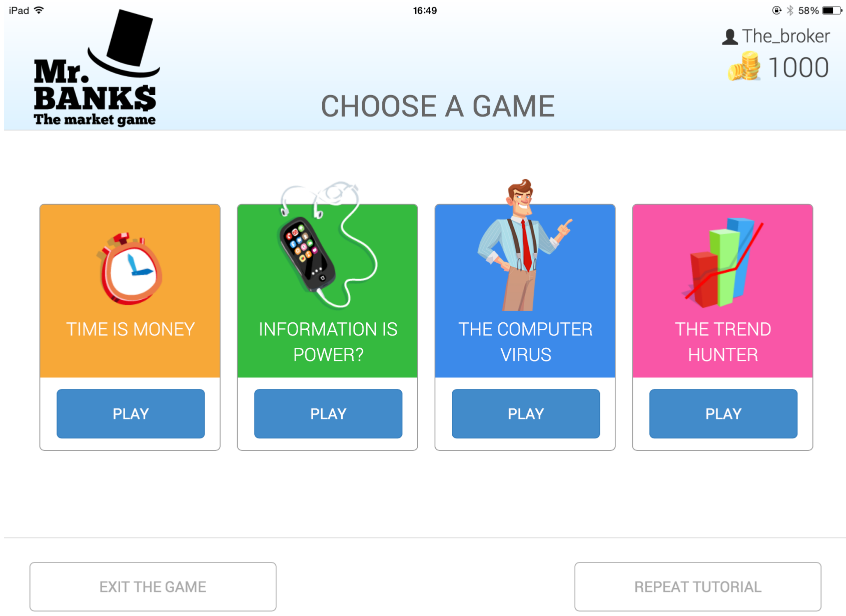

Figure 4.1: Mr.Banks games. Home screen presenting the four games: (1) Time is money, (2) Information is power?, (3) The computer virus and (4) The trend hunter.

These two experiments are carried out in real environments, outside the laboratory, which is not usual in experimentation. To this end, citizen science techniques are applied, placing citizens at the center of scientific research. With these two experiments we built the basis for the experimentation platform, which can carry out experiments with a light infrastructure and with gamified elements to engage the participants, because in the field they are not recruited beforehand. So much so, that the first decision-making experiment, in the context of the stock-market, is called Mr.Banks and the second, in which the participants play the dyadic games, is called Dr. Brain.

The data collected in the experiments allows for the analysis at a group and individual level. The participants are characterized by their sociodemographic data and by a set of decisions. Specifically in Dr. Brain, the actions of the participants are collected in each of the four games and within each game in each discretized game point. In order to analyze the individual behavior of all participants among an unstructured set of decisions, the clustering techniques can be useful to naturally give rise to patterns in the case that they exist, and thus we are able to define individual behaviors in the participants.

In this chapter we present the first two experiments that were carried out with the platform that we don't discuss in detail because them are included in another thesis (Gutierrez-Roig, 2015). Instead, in the first, Mr. Banks, we evaluated the subject pool and the results between the main experiment and the two replicas, which allow to shed light on the quality of the data collected by the platform in different environments. In the second, Dr. Brain, we focus on the techniques of data analysis to find patterns in unstructured data and some methods to check the robustness of the algorithms. Finally, we characterize the behaviors and describe the phenotypes that emerge from the analysis, as well as the outcomes that this type of experiments have beyond the pure scientific impact.

4.2 Mr.Banks

Mr. Banks presents the stock market as an experimental scenario in which the participants are in the position of a financial trader, analyst or, in general, an investor that must predict, based on multiple pieces of information, whether the following market movement will be up or down.

With this in mind, at the beginning of the experiment a series of stock data taken from real historical records of markets [1] is presented randomly to each participant. The information is displayed in five screens. The first and main screen contains the daily price movements from the round and during the previous 30 days from that round, as well as a 5-day moving average and/or the 30-day moving average. The second screen shows a chart of the intraday price of the day before. The third screen shows an expert's advice, affirming the current volatility of the market (low/high) and a possible price tend (up/down). The possibility of the expert's advice being correct was of 60%, but participants were not aware of this. A fourth screen shows a simplified market evolution, only the directions (ups and downs) from market data of the last 30 days. Finally, a fifth screen includes information of nine other indexes[2] from three different continents of the last three days.

Participants have to decide whether the market will go up or down. For each participant we register 100 decisions spread over 4 different scenarios and in 25 rounds for each scenario. They have 30 seconds to make a decision. The participants have restrictions in each of the four scenarios. (1) "Time is money": all the participants have all the information available, however 50% of the participants have 30 seconds to decide while the rest have 10 seconds. (2) "Information is power": 50% of the participants have all the information, while the rest can only access the main screen with the daily prices from that round and from the previous 30 days. (3) "The computer virus": participants can only access the main screen and another screen, which in 50% of cases is already chosen for them, and in the 50% remaining cases can be chosen by the participants. (4) "The trend hunter": 50% of the participants can see all the information available, while the rest can see the market directions, therefore they know the trend (bearish, bullish or flat).

As mentioned earlier, the platform presents the information distributed on different screens according to the decision-making scenario in which the participant is placed. In Mr. Banks, the interaction of the participants is individual, that is, they make decisions based on the series that they have played in a random way without interacting with other participants. The platform is fully gamified to engage participants and, therefore, they continue making decisions, since each scenario has 25, which in total add up to 100, so it is necessary to encourage participation through game dynamics. To do this, a Mr.Banks character is created and the game is aesthetically pleasing. The participants start with a total of 1000 coins and, as the game progresses, they increase or decrease their points by 5% depending on whether or not they make good predictions.

4.2.1 Experiment replicability

Mr.Banks was developed to be carried out during DAU Festival, a board game fair held in Barcelona during the weekend of the 14th and 15th of December 2013. This is the main experiment in which 283 participants took 24503 actions, and the results are published here (Gutiérrez-Roig et al., 2016b). Besides the main experiment performed at DAU Festival, two

1. Concretely, the data taken are 30 series of 25 consecutive days of stock data picked from 01/02/2006 - 12/29/2009 of daily prices of: the Spanish IBEX, the German DAX and the S&P500 from the United States. The 30 series show different tendencies, specifically 10 with a downwards trend (bearish market), 10 with upwards trend (bullish market) and 10 with no specific trend at all (flat market).
2. S&P500, DJI, NASDAQ, FTSE, IBEX, CAC, DAX, NIKKEI, HSI.

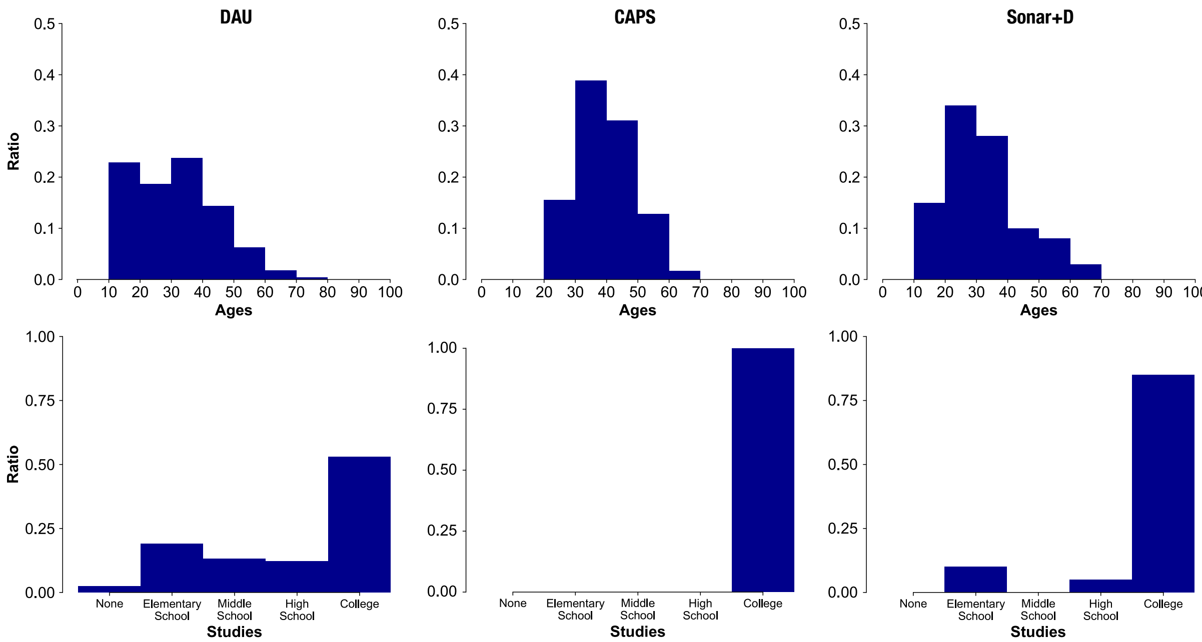

Figure 4.2: Sociodemographic differences in the experiment replication. Differences in the distribution of age and educational level in the experiments of DAU (n=283), CAPS (n=37) and Sonar+D (n=20).

replicas of it were conducted. One in Sonar+D, a music technology fair held in Barcelona in the context of Sonar, the electronic music festival, on June 6th, 2015, in which 20 participants contributed with 1,078 actions. The other took place at CAPS 2015, the Collective Awareness Platforms for Sustainability and Social Innovation conference held in Brussels during the 7 and 8 of July 2015, where 37 participants contributed with 2397 actions.

The experiments were carried out in a controlled area and they were assisted for researchers the whole time in order to answer any questions related with the experiment. Here, we present the way in which, thanks to the versatility of the platform, we are able to perform experiments following the pop-up experiment principles (Sagarra et al., 2016) in different locations and achieving robust results independently of the conditions of the experiment.

Participation

The participation in the main experiment at DAU was much higher than in the two replicas. There were also differences in sociodemographics. Regarding gender, the proportions were similar, and only between 30-35% were women. With regard to ages, at the DAU and Sonar+D experiments the participants were between 16 and 35 years old, although at the DAU, unlike Sonar+D, approximately 30% of the participation was under 15 years old. At CAPS, however, most participants were between 26 and 55 years old. With regard to the level of studies, at the replicas of the Sonar+D and CASP between 85% and 100% of the participants were university graduates. Although, in the case of the DAU experiment, the majority (52.30%) had university studies, and the rest was divided between Elementary and High school.

The experiments were performed in different environments, and therefore the audiences were also different. In addition, the DAU experiment was carried out with a very clear objective of collecting as many samples as possible in order to have enough statistics to study the emergence of strategies and behaviors. However, the replicas were more focused on expanding the sample and checking the robustness of the platform. We can see how the sample of participants is very

Table 4.1: Mr.Banks participation and actions. Probability to decide "Up" and "Down"

Experiment	Participants	Men	Woman	Actions	Up	Down
DAU	283	184	99	18436	0.61	0.39
CAPS	36	24	12	1895	0.61	0.39
Sonar+D	20	14	6	777	0.63	0.37

diverse, so these experiments, depending on the environment where they are developed, allow for the recruitment of very different participants in sociodemographic terms.

The participation presented differences between the DAU main experiment and the two replicas of the Sonar+D and CAPS. At the DAU experiment, the majority of participants (77.39%) engaged in all four scenarios, while in the other experiments between 40-50% participated in all four scenarios and practically the rest only tested one scenario.

The main results of the three experiments are shown below in order to prove not only the replicability of the experiment itself, but the robustness of the platform to collect quality data in different contexts. To illustrate those results, we present the outcomes of the DAU experiment and the two replicas. First, the number of "up" and "down" actions aggregated in each context (Table 4.1). Next, the strategies adopted by the participants in each case, which are Market Imitation (Heyes, 2011) and Win-Stay Lose-Shift (Wang et al., 2014; Nowak and Sigmund, 1993). Market Imitation consists on replicating the input received in response to the market instead of waiting for the outcome (Table 4.2). On the other hand, Win-Stay Lose-Shift strategy consists on repeating the last decision if it was correct and changing whenever it was wrong (Table 4.3). The experimental results show how, even with a limited number of samples, in the experiments' replicas the data collected is robust. In fact, only two cases show significant differences. First, between DAU and Sonar+D experiments there are differences (-2.53 SD) in the probability to select "Up" whether the market goes Up, and second, between DAU and CAPS experiments there are differences (2.35 SD) in the probability to change the selection whether the previous selection was wrong. These may be due to differences in the behaviour

Table 4.2: Market imitation. Biases with respect to the market (Participant/Market)

Experiment	Up/Up	Up/Down	Down/Up	Down/Down
DAU	0.71	0.29	0.47	0.53
CAPS	0.71	0.29	0.49	0.51
Sonar+D	0.78⋆	0.22	0.46	0.54

⋆ There are significant differences (-2.53 SD) between DAU and Sonar+D experiments (Binomial process differences test).

Table 4.3: Win-Stay Lose-Shift strategy. Decision conditioned to performance (Strategy/Decision)

Experiment	Win/Stay	Win/Shift	Lose/Stay	Lose/Shift
DAU	0.68	0.32	0.42	0.58
CAPS	0.71	0.29	0.48	0.52⋆
Sonar+D	0.72	0.28	0.40	0.60

⋆ There are significant differences (2.35 SD) between DAU and CAPS experiments (Binomial process differences test).

of participants in each experiment or a lack of data, in any case the results of the replicas are coherent with the main experiment.

4.3 Dr.Brain

Dr. Brain presents a set of dyadic games in which the participants interact with each other making decisions ("cooperate" or "defect") in different scenarios that imply different consequences, because the tension between these decisions varies in each interaction.

In this experiment we use four simple dyadic games which cover the TS plane and capture two important features of social interactions: the temptation to free-ride and the risk of cooperation. The games used are the prisoner's dilemma, stag hunt, snowdrift and harmony game. In each of the games the tension between cooperating and defecting varies. So much so that in Harmony, the individual and collective benefit match, so there is no tension, while in Prisoner's dilemma it is just the opposite, and the tension is maximum. Within each of the quadrants of the TS plane in which the games are located there are different tension equilibria. To cover several possibilities, the TS plane is discretized as a matrix of 11 x 11 so that in total we have 121 games with different tensions.

The experiment was carried out in groups of a maximum of 25 people playing simultaneously. Each of them was randomly paired with another participant, and each pair was assigned a point on the plane also randomly. The pair of participants took the decisions of cooperating or defecting and obtained the outcome as the result of both . This was repeated between 13 and 18 rounds, and in each round the point of the plane and the pair of participants interacting changed. Therefore, the participants were not informed of the number of rounds they played, nor of the participant with whom they interacted. In addition, the matrix changed in each round, so each decision-making was different from the previous one.

The Dr.Brain experiment was performed at the DAU Festival, held in Barcelona during the weekend of the 13th and 14th of December 2014 following the practices of pop-up experiments. In total, we recruited 541 participants of different ages, educational levels and social status. The participants made decisions during multiple rounds in which participant were randomly assigned partners and games (payoff values), allowing us to study the behaviour in the four dyadic games mentioned. The participants were shown a brief tutorial with detailed information (see Appendix A) and the complete results of the experiment are published in (Poncela-Casasnovas et al., 2016).

4.3.1 The experimental platform

To conduct the experiment and collect the data we created a new setup using the experimental plaform. In this case we performed the experiment using a local network architecture which consisted in 25 mobile devices (tablets), a router, and a laptop running a web server and a database server. The system was designed to allow playing synchronized sessions, to collect and store user data safely, and to control the experiment in real time while the users were playing against each other. The game was accessible through a web application specifically designed for tablets. All the interactions that users made through the game interface were immediately sent to the server through a client API. The server also provided a server API to control and monitor the status of each experiment session.

4.3.2 Data Analysis

The data collected by the platform correspond to the decisions,"cooperate" or "defect", that the participants took in each round. Each round represents a different TS-plane position, so, during the whole game, the participants made decisions in different games of different tensions. In the Fig.4.3, the total number of actions in each point of the TS-plane is represented for all 541 participants in the experiment, and the total number of game actions in the experiment adds up to 8366. In average, around 69 for each point.

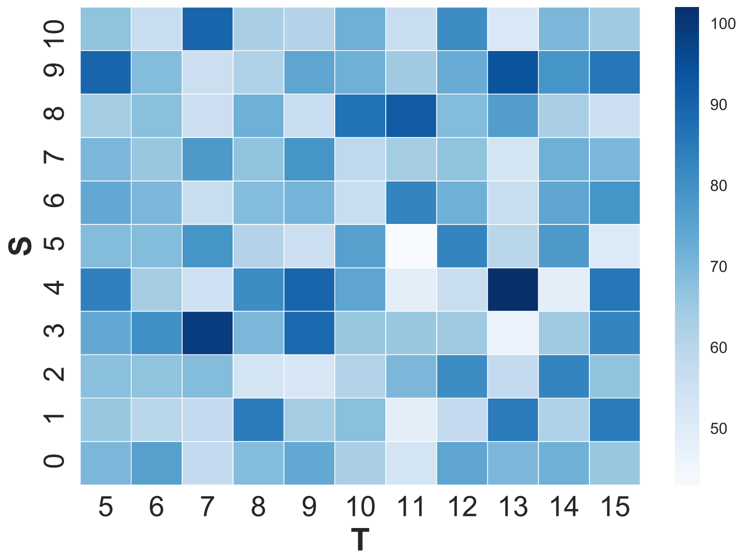

Figure 4.3: Total number of actions in each point of the TS-plane. For all 541 participants in the experiment the total number of game actions in the experiment adds up to 8366.

We want to take a close look and try to understand the behaviours at an individual level. For each of the participants, we build a vector with all the decisions that they take on the TS-plane. The decisions that each participant takes fall in one of the four quadrants, therefore in each of the four dyadic games. We analyze the individual behavior in each of the games and observe whether any behavior pattern emerges among the participants' decisions or whether, on the other hand, they behave in an distinctive way. This way, we create a vector of four values, each of them representing the average cooperation in each of the games. Essentially, we represent each participant with a four-dimensional vector where each value represents the proportion of cooperation in each of the games.

Clustering and Phenotypes

With the goal of looking for behavioral patterns, we adopted machine learning techniques to find groups of users who behave in a similar way. Specifically, we used unsupervised learning techniques, since we do not have any previous structure in the collected data, and consequently applied clustering algorithms to our dataset, not pre-setting any criteria. We applied the k-means clustering algorithm to group individuals with similar cooperation vectors.

The k-means algorithm groups the data in a number of clusters that are given by the data analyst. Therefore, it is first necessary to find the structure of our dataset. To find out the

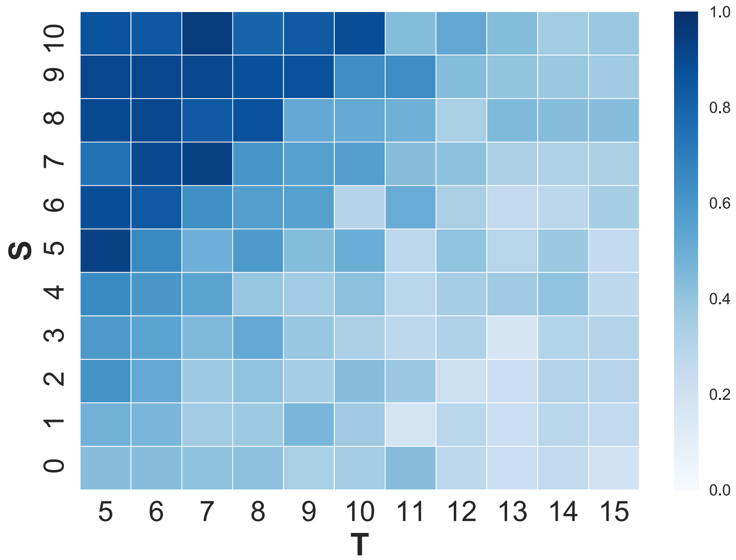

Figure 4.4: Summary of cooperation in the games. Average empirical cooperation from the 8366 game actions of the 541 participants, in each cell of the TS-plane.

number of groups that fit well within the dataset we used the Davies-Bouldin index. We calculated 200 times the index per k = [2,20], with different seeds in each of the interactions, and then calculated the average of all of them. The result is that the ideal number of groups in the database is k = 5, all without having set the number of behaviors in advance. By applying the k-means algorithm without any precondition we got five groups. By their levels of cooperation, we can say that there is a group that mostly cooperates in the harmony game, a second group that cooperates in both the harmony game and snowdrift , and a third one that cooperates in both harmony game and stag-hunt. Players in the fourth group cooperate in all games, and, finally, we find a small group who seems to randomly cooperate almost everywhere.

The group of optimist (20% of the population) and pessimist (22% of the population) display different behaviors in coordination (cooperating in stag-hunt game) and anti-coordination (cooperating in snowdrift game) respectively, although they both act according to the Nash equilibrium in prisoner's dilemma and harmony game. The optimists maximize the maximum payoff and the pessimists maximize the minimum payoff. The envious (30% of the population) exclusively cooperate in the upper triangle of harmony game. In the same vein as with optimists and pessimists, this third behavior is not rational, in so far as players renounce the possibility of achieving the maximum payoff choosing the Nash equilibrium in harmony game. These players basically attempt to ensure that they receive more payoff than their opponents. The trustful (17% of the population) cooperate in almost every round and in almost every site of the TS-plane. In this case, and opposite to the previous one, we believe that these players' behavior can be associated with a trust in their partners behaving in a cooperative manner. Last, the unsupervised algorithm found a small fifth group of players (12% of the population) who cooperate randomly, with a probability of 0.5, in any situation, we call this group as the "undefined".

Robustness of phenotypes

Lacking the ground truth behind our data in terms of different types of individual behaviours, we test the robustness of the phenotypes, which is to say the robustness of the clusters, in

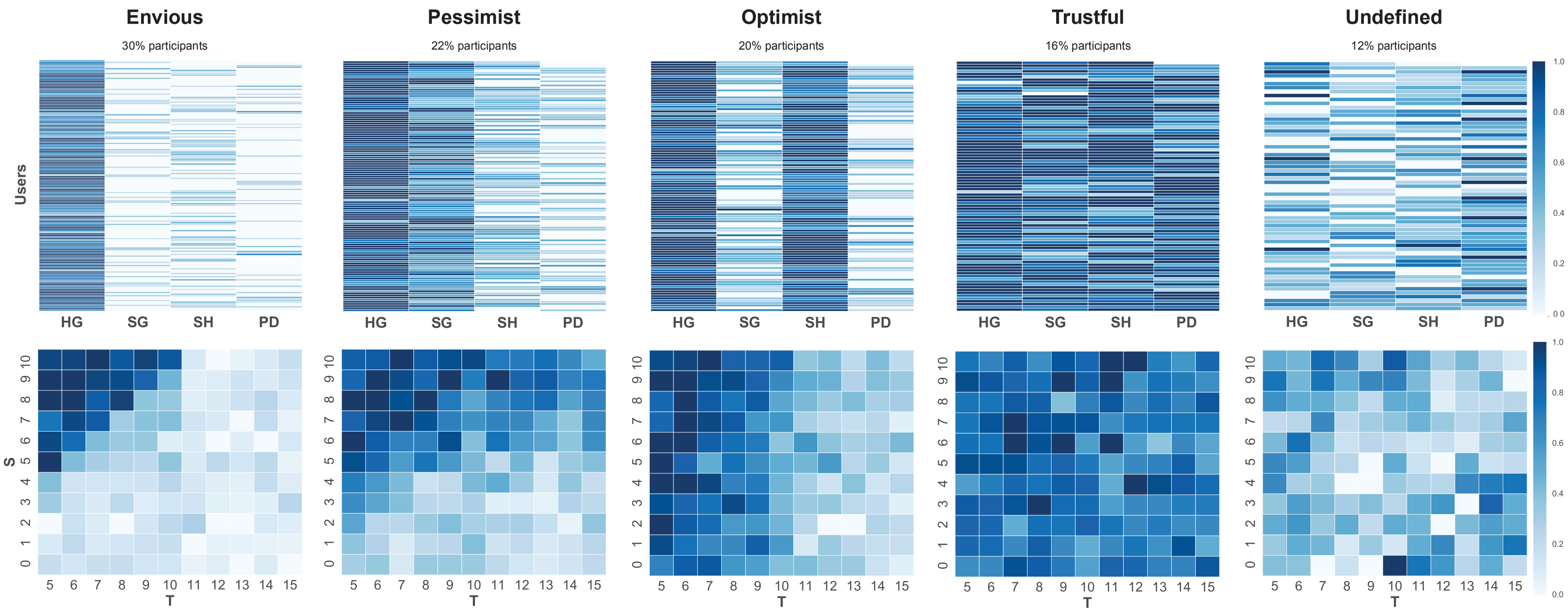

Figure 4.5: Summary of behavioural phenotypes. Running the k-means algorithm emerge the behavioural phenotypes (envious, pessimist, optimist, trustful and undefined) represented in here from the decisions of participants in each TS-point. The top row shows the levels of cooperation per participant in each game, we can observe patterns in each of the phenotypes. In the row below, the level of cooperation is represented in each point of the TS-plane, observing, again, cooperation patters.

several complementary manners.

We randomize our data and apply the same algorithm, so we preserve all cooperative actions but associate them with random users, thus eliminating the action-participant correlation. We take the 8366 actions of the 541 subjects and create a pool with them. From this pool of data we replace the real decisions to obtain new randomized sets of actions for each person. In such a way we preserve the number of times each participant has played and the particular TS-points and the average fraction of cooperative actions in the population, but destroy any possible correlations among the actions of any given subject. We observe the lack of internal structure of the randomized data.

We apply the leave-p-out procedure (Kohavi, 1995), running the k-means algorithm with portions of our data, to test our clusters against data perturbations. This test showed that the optimum five-cluster scheme found is robust even when randomly excluding up to 55% of the players and their actions. In order to do so, we run the algorithm 200 times again, but each time we exclude a randomly chosen number of players and all of their actions. We do this for a scheme with k = [2,20] clusters, and leaving out p = 100, 300, 400, and 450 subjects (out of the total 541), and calculate again the average DB index for them.

We tested the consistency among cluster structures found in different runs of the same algorithm for a fixed number of clusters. To ascertain this, we computed the normalized mutual information score MI (MacKay, 2005), knowing that the comparison of two runs with exactly the same clustering composition would give a value MI = 1 (perfect correlation), and MI = 0 would correspond to a total lack of correlation between them. We ran the k-means clustering algorithm 2000 times for the optimum k = 5 clusters and paired the clustering schemes for comparison, obtaining an average normalized mutual information score of MI=0.97 (SD: 0.03).

4.4 Discussion

Mr.Banks represents a decision-making experiment with a very simple interaction, since the participant takes decisions independently of the other participants. However, Mr. Banks helped us to test the platform in terms of the participant-centered design with elements of gamification, experimentation in environments outside the laboratory with light infrastructures and the collection of data in a robust manner.

The results presented in Mr.Banks, beyond the value in the field of human behavior and decision making, show the feasibility of conducting experiments in environments outside the laboratory, with a diverse sample pool depending on the context where it is carried out and with the possibility of mounting the infrastructure in a very simple way and obtaining quality data even with small samples.

On the other hand, Dr.Brain, the second experiment that we carried out included a more complex interaction system. In it, the participants interact with each other and the outcome is a consequence of the decisions of both participants. The execution of the experiment during two days allowed a very high participation and with very different sociodemographic data.

If we focus on the analysis of the data obtained from both experiments we can see how unsupervised learning techniques can be very useful to find behavior patterns in data that has not been previously classified, nor structured. Given that, this structure does not exist, and therefore it is not possible to corroborate the accuracy of the results, so it is necessary to find mechanisms that allow the calculation of the results' robustness.

Focusing on the results of the experiment in terms of behavior, we see how a set of patterns emerge from the decisions that have been made in the different games, which in the terminology of Peysakhovich et al. (2014) are called phenotypes. The population that participated in our experiment can be classified into a reduced set of five phenotypes: envious, optimist, pessimist, trustful and a small group of individuals with an undefined behavior.

The interesting thing about the groups that form is that they emerge directly from the data without any presumption thanks to the unsupervised learning algorithm and whose robustness is validated from different approaches. In addition, the phenotypes presented here can be interpreted from different theories that have tried to classify individual behaviors, such as the social value orientation (Balliet et al., 2009; Rusbult and Van Lange, 2003; Van Lange, 2000) or social preferences theory (Fehr and Schmidt, 1999).

The impact of the experiments' results has direct social implications in policy-making, and these results are useful for organizations that want to create more efficient collaborative environments, since they describe interactions that occur in everyday life. Similarly, it should be noted that the results show evidence of the utility of the behavioral suite of games to better understand people's behavior, particularly in experimentation outside the laboratory, where it is proven that very different behaviors emerge. Regardless of the impact of determining behavior at the individual level, it is essential to highlight the importance of using the human behavior experiment approach in real environments, outside the laboratory, focusing on social decision-making and impacting on the policies (The Behavioural Insights Team, 2018).

The first two lab-in-the-field experiments that were carried out with the platform of human behavior, and that are detailed here, allowed to lay the foundations of all the behavioural experiments that have been carried out afterwards. The experiments have been successful in terms of participation, data quality, user experience, and they have generated scientific outcomes.

Chapter 5

Collective Behaviour and Inequalities in the Fight Against Climate Change

Summary – Climate change mitigation is a shared global challenge that involves collective action of sets of individuals with different tendencies to cooperation. However, we lack an understanding of the effects of resource inequality when diverse actors interact together towards a common goal. Here, we report the results of a collective-risk dilemma experiment in which groups of individuals were initially given either equal or unequal endowments. We found that the effort distribution was highly inequitable, with participants with fewer resources contributing significantly more to the public goods than the richer −sometimes twice as much. An unsupervised learning algorithm classified the subjects according to their individual behavior, finding the poorest participants within two "generous clusters" and the richest into a "greedy cluster". Our results suggest that policies would benefit from educating about fairness and reinforcing climate justice actions addressed at vulnerable people instead of focusing on understanding generic or global climate consequences.

5.1 Introduction

Mitigating anthropogenic climate change (Pachauri and Meyer, 2014) is a complex problem involving many heterogeneous actors with different agendas and conditions (Dennig et al., 2015; Raupach et al., 2014; Mendelsohn et al., 2006; Azar and Sterner, 1996).

While climate change mitigation requires a collective action (Olson, 1971), it is not clear what are the effects of the inherent diversity of the agents involved in it (Ruttan, 2006). There is the risk that the poor exploits the rich, i.e., that the largest beneficiaries of a common goods bear a disproportionately large burden in its production (Olson, 1971). Conversely, the poor also have an incentive to contribute since they are risking as much as the rich in the event of a catastrophic development.

Inequality between actors is a widespread situation related to climate change mitigation which has an important impact within a given urban and shared space. For instance, increasing green spaces and similar actions that improve urban environmental health are part of the recipes within the Paris Agreement (COP21) (Rhodes, 2016). However, these actions impact differently more and less affluent citizens and, in fact, such actions have also been questioned from the perspective of environmental justice (Wolch et al., 2014). This is the reason for the increasing interest in linking climate justice with socio-economic inequalities and in incorporating some behavioral aspects in city policy (Bulkeley et al., 2014). Therefore, understanding how to deal with climate change mitigation in an economically diverse world and in an environmentally fair manner is of special interest. This has become more timely and pressing with the withdrawal (Trump, 2017) of the second largest CO_2 emitter (PBL Netherlands Environmental Assessment Agency, 2016) from COP21.

Resource heterogeneity and environmental justice can be very suitably framed within the collective-risk dilemma experimental setup introduced by Milinski et al. (2008) a decade ago. In this framework, to be described in detail below, groups of people had to reach a common goal by making contributions from an initial endowment. If the goal was reached, every individual received a gift card containing the part of the money not contributed. If not, a catastrophe occurred with certain probability, and all participants lost all the money they had kept. While many experimental and theoretical studies have considered different aspects of climate change within this framework (Dutta and Radner, 2009; Milinski et al., 2011; Tavoni et al., 2011; Abou Chakra and Traulsen, 2012; Santos et al., 2012; Burton-Chellew et al., 2013; Hilbe et al., 2013; Jacquet et al., 2013; Abou Chakra and Traulsen, 2014; Freytag et al., 2014; Vasconcelos et al., 2014; Dannenberg et al., 2015; Bynum et al., 2016; Hagel et al., 2016; Milinski et al., 2016; Hagel et al., 2017), the issue of heterogeneity has only been considered in three experiments framed in the climate change issue. Thus, Tavoni et al. (2011) included inequality by separating the participants into two groups with different starting endowments, and found that the common goal was less likely to be reached. However, when they allowed participants to communicate their intentions, the probability of reaching the target goal increased again, similarly to conditional cooperation in public goods games (Fischbacher et al., 2001). On the other hand, Milinski et al. (2011) observed that when they included rich and poor subjects, rich ones substituted for missing contributions by the poor provided intermediate climate targets. However, despite this increase in the contributions of the rich, the final target was reached less often than the intermediate target. Finally, Burton-Chellew et al. (2013) proposed a game with four conditions varying the initial endowments and/or the risk of a catastrophic climate event. They found that inequality in both endowments and risk decreased cooperation, that is to say, selfishness emerged when rich were less at risk. However, some rich players were still reluctant to cooperate when they suffered the higher risk. It seems that climate change awareness could have mediated their responses, since it was found to be proportional with individual contributions. Outside of the climate change framing, studies on the role of heterogeneous endowments in public goods games are not conclusive, with different experiments leading to different results (Chan et al., 1999; Cherry et al., 2005). Therefore, new studies are needed to understand the influence of heterogeneity in threshold public goods games.

Here, we largely extend the knowledge on the effects of resource heterogeneity in collective dilemma in two main directions. First, we include broader capital distributions, thus representing more closely the diversity in resource availability among the members of a given collective public goods dilemma, e.g., between different countries worldwide or inhabitants in a given urban context. Second, and most importantly, we go beyond aggregate results to analyze the behavior of individuals by means of agnostic classification tools that allow us to identify differences between the behavior of subjects with the same resources. We complement this with a questionnaire probing into the subjects' knowledge of the climate change crisis and the influence of such knowledge in their actions. Therefore, we provide a much more complete picture of the influence of inequality that encompasses both the global (reaching the collective goal) and the local (how different individuals behave under different circumstances) visions of the problem. Such a two-level approach is the best option in order to identify how agents react to resource heterogeneity and what actions must be taken to promote environmental justice. As we show below, our findings allow to hint directions for policy measures targeted to specific collectives.

5.2 The Collective-Risk Dilemma

5.2.1 Game definition

The original collective-risk dilemma (Milinski et al., 2008) introduced groups of six people where each person receives an initial capital (40 €) and the common goal of the group is to collect 120 € that will be invested in mitigating climate change (by publishing an ad in a national newspaper). The game consists of 10 rounds, and at every round each subject decides how much she contributes to the common fund (0, 2, or 4 €). If the goal is reached at (or before) the end of the game, all participants keep the money that they have not contributed. Otherwise, the participants only keep the remaining money with a probability which in (Milinski et al., 2008) was 90%, 50%, or 10% (equivalently, a climatic catastrophe occurred with probability 10%, 50%, or 90%). In addition, in this case no money goes to climate change mitigation. The main result was that most groups did not reach the goal, and even in the worst case scenario (catastrophe probability 90%) only about half of the groups avoided climate change. We chose this worst case scenario as our baseline treatment, and we carried out experiments with the equal distribution for comparison with the unequal one.

To introduce inequality, we distributed six different windfall starting capitals (20, 30, 40, 40, 50, and 60 €) randomly amongst the participants. In half of the games the participants could invest 0, 2, or 4 € per round as in the above setup, while in the other half of the games we allowed them more flexible choices, namely 0, 1, 2, 3, and 4 € per round. In all cases we informed the participants that in case they reached the goal of 120 €, the so collected money would be used for a reforestation action by planting trees in a nearby park with the help of an NGO organisation (NGO, 2016). Finally, after the experiment, we asked our subjects to answer a questionnaire (see Appendix B.4) to have an individual assessment of climate change awareness and predisposition to collaborate in common actions that could be further correlated with the individual's contributions.

5.2.2 Equilibria and fair distribution

In our heterogeneous version of the collective-risk dilemma, the unequal treatment, there are very many Nash equilibria (Nash, 1950), which makes claiming that one or other behavior should be expected very difficult. Indeed, in the equal case, the number of equilibria can be refined by imposing symmetry (meaning that all subjects, being equal, should choose the same contribution and, therefore, obtain the same payoff). Then, two equilibria are left, with individuals either contributing nothing or contributing exactly 20 €. Of course, these equilibria refers only to accumulated contribution along the game; considering the different sequences of investments in the 10 rounds recovers the multiplicity of equilibria and we will not consider them.

In our unequal setup we have 5 types of players (there are two endowed with 40 €) and the symmetry refinement no longer holds. Subjects contributing nothing is still a Nash equilibrium that leads to expected gains of 10% of every subject's endowment. It is then easy to show that any combination of individual total investments that adds up to exactly 120 € such that every player makes more money than in the "contribute nothing" case is also an equilibrium. Therefore, there is not a clear theoretical prediction about what should happen in our heterogeneous version of the game.

While in the equal case in which the fair share behaviour was given by equal contribution and equal payoff (50% of the endowment in both cases), the asymmetric capital distribution in the unequal scenario requires three different concepts of fairness based on the same ideas. The first one refers to the equal payoff, "payoff fairness" (Fig. 5.1b); the second one considers the criterion of equal contribution, "contribution fairness" (Fig. 5.1a); and, finally we also analyze a "relative payoff" based on contributing half of the endowment (Fig. 5.1c). The definitions of these quantifications of fairness are as follows: "Payoff fairness" considers as fair behaviour the one in which each participant contributes an amount resulting in the same final payoff for all (i.e., 20 €). This means that participants receiving an endowment of 20 € should contribute 0 €, participants with an endowment of 30 € should contribute 10 €, and so on. "Contribution fairness" defines as fair an equal contribution (20 €) of each participant independently of the endowment. Finally, "relative payoff" considers as fair a 50% contribution of every subject's initial endowment (i.e., 10 € for participants starting with 20 €, 15 € for participants starting with 30 €, and so on.)

Among these possibilities, we decided to take as the main reference for the discussion of our results the "relative fairness", which is the most representative of the inequalities generated by the participation in the game. If the participants contribute following "payoff fairness", they start from a position of inequality (participants endowments from 20€ to 60€) and once the game finishes their payoff is the same for all of them (20 €), breaking and unbalancing the initial inequalities due to the participation in the game. On the other hand, "contribution fairness" does not consider endowments at all, and therefore the initial inequality. However, in the case of "relative fairness", after participating in the game the subjects maintain the same inequality distribution they started with. Therefore, we believe that this definition allows us to analyze the impact of the collective action in a more equitable and proportional way.

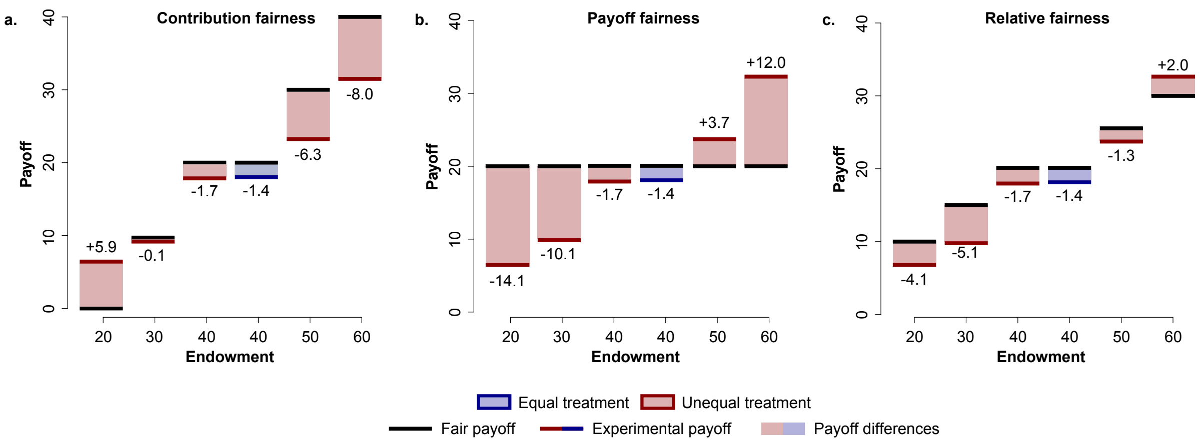

Figure 5.1: Fairness definitions payoff and experimental payoff in both equal and unequal treatments. The values in the graph represent the differences between the experimental payoff and the fair payoff for each fairness definition and endowment. **a. Contribution fairness:** payoff of the fair behaviour based on equal contributions. This definition yields particularly negative results for the most vulnerable participants, those with 20€, that do not get any reward at all (and therefore would prefer to keep their endowment so they have a chance to earn something). **b. Payoff fairness:** payoff of the fair behaviour based on equal payoff. The differences in the experimental payoff increases because it forces the participants to end up with the same reward. **c. Relative fairness:** payoff of the fair behaviour based on contributions of half of initial endowment, this definition maintains the same inequalities as before participation.

5.3 Results

5.3.1 Collective climate action

In all games played in our experiment, the participants reached the goal irrespective of the initial endowments being equal or unequal. In the former case, our result has to be compared with only 50% success rate for the groups in (Milinski et al., 2008). An increase in groups reaching the target has been observed in other similar studies carried out later (Milinski et al., 2011). The evolution of the games also differ from previous experiments (see Fig. 6.4): In all the treatments in our experiment, the sum of money accumulated at the end of each round is always above the fair contribution (12 € per round), that is, participants contribute much more in the initial rounds, making the group reach the goal faster, and then they stop contributing at the end once the goal has been secured (see details in Figs. 5.7 and 5.9). In contrast, the original results in (Milinski et al., 2008) showed contributions below the fair one until the end of the game, and those groups that reached the goal did it by increasing their contributions in the last rounds.

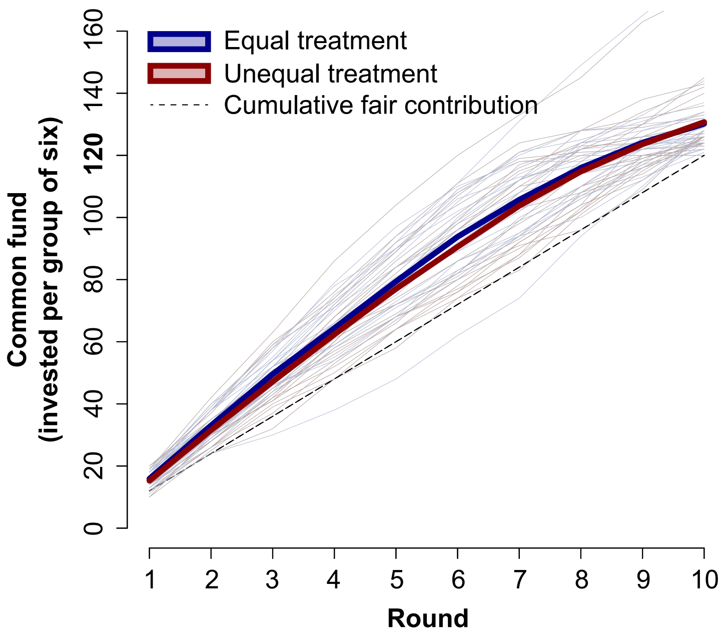

Figure 5.2: Average contribution to common fund per round in the Climate Game for both, equal and unequal treatments. Equal treatment consist of 24 valid games in which all players are endowed with 40 €), and unequal treatment (endowments are 20, 30, 40, 40, 50, and 60 €) with a total of 26 valid games. Both treatments show an accumulated contribution over the game evolution above the fair contribution per round.

5.3.2 Effect of unequal capital distribution

Figure 5.3 presents the average amount of capital contributed as a function of the initial capital of the participants. We observe that, in terms of absolute contribution, the subjects with high endowments, 50 and 60 €, are the ones that contribute the most, 2.6 ± 0.13 € and 2.8 ± 0.12 € per round respectively (mean±SE). Participants with low-endowments, 20 and 30 €, contribute

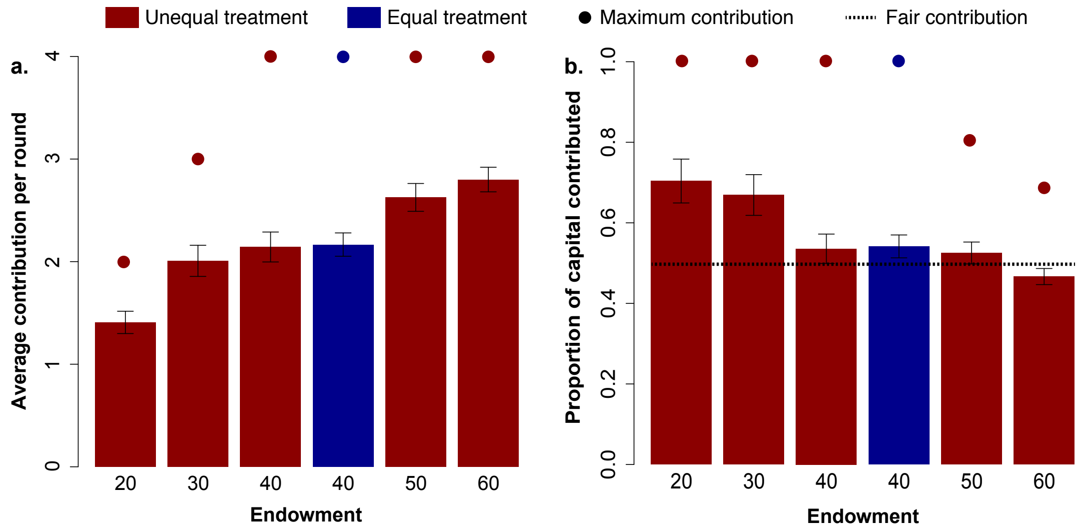

Figure 5.3: a. Average (95% CI) capital contributed according to the participants endowments in both treatments, equal and unequal. Note that participants starting with 20 € and 30 € can only reach a maximum average contribution per round of 2 € and 3 € respectively. **b. Average (95% CI) proportion of capital contributed according to the participants endowments in both treatments, equal and unequal.** Dotted line represents the fair contribution, which we have defined as contributing 50% of the initial capital. The effort to contribute is different depending on the endowments, so dots represent the maximum investment that each group can reach. Participants with endowments of 50 and 60 € always keep a proportion of capital as savings even if they contribute the maximum amount of 4€ per round.

the least,1.4 ± 0.11 € and 2 ± 0.15 € respectively. However, this comparison is not the best for interpreting the obtained results, since the initial capital of the poorest players only allows them to contribute a maximum of 2 € per round (earning 0 at the end). Therefore, the comparison makes more sense in terms of the percentage of capital contributed relative to their total capital, which in turn allows to discuss the results using the "relative fairness" as reference (see Fig. 5.4). Strikingly, we observe that the most affluent (endowments of 60 €) are the ones that contribute proportionally less, with around 46.6% of their initial capital, while the poorest (starting with 20 €) contribute around 71.4% of their initial capital which shows their vulnerability when facing the collective risk dilemma. Figure 5.3 shows very clearly the stark contrast between the two visions. To put this result further in context, we notice that the maximum contribution from participants with a starting capital of 20 € (2 € per round) implies an effort of 2 times the fair contribution, whereas for participants with a starting capital of 60 €, the effort of contributing 2 € per round is 0.66 times the fair contribution. Therefore, in that case, contributing 2 € have unequal impacts in the endowments of subjects: poor people, contributing 1.4 € on average, make a larger effort than that of the rich contributing 2.8 €, which is even below their fair share of the threshold.

5.3.3 Individual behaviors

Once we have looked at the average evolution of the different groups, we focus on the individual behavior and study whether participants with the same starting capital behave similarly. We characterize the set of decisions taken by each participant with a vector, grouping decisions by

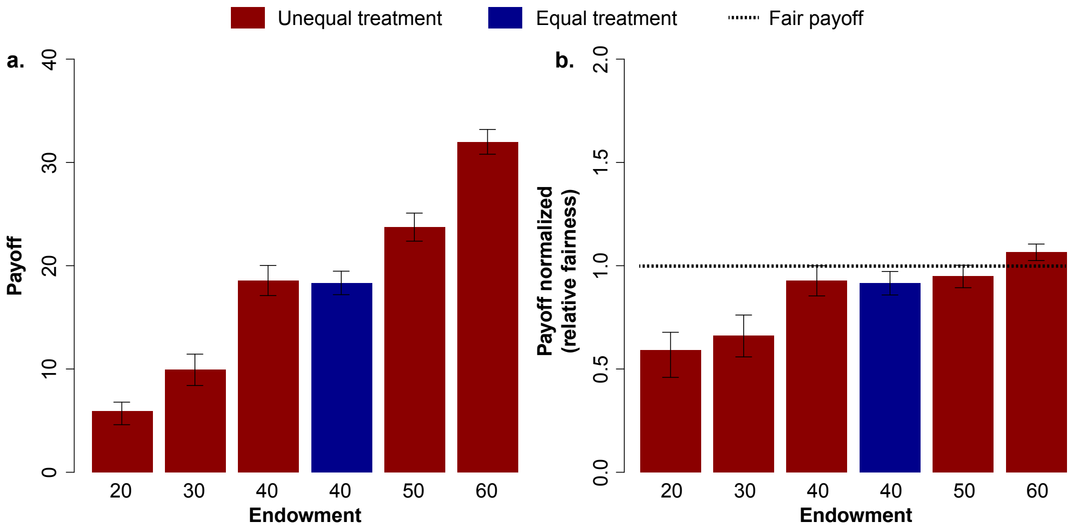

Figure 5.4: a. Payoff and b. Payoff normalized. Average (95% CI) payoff, it is normalized following the definition of "relative fairness" that allow us to study how the rewards have been distributed. Those who had more resources get high payoff, even above the fair payoff. The most vulnerable (20€ and 30€) get low payoff and there are significant differences between them and the rest of participants with high endowments (see Appendix C.1, Table C.1 and Table C.2).

the capital on the common fund at the beginning of the round (see Table 5.1). This is motivated by the intuition that subjects choose their contributions as a function of their endowment but also taking into account the current situation and whether the goal is closer or farther. In turn, we can monitor how the contributions differ depending on the stage of the game where they are (see Figs. 5.11 and 5.10). Next, in order to detect, identify and characterize different types of behavioral patterns, we use an unsupervised learning algorithm, namely Ward's hierarchical clustering method (Ward, 1963; Murtagh and Legendre, 2011) with squared Euclidean distances. Additionally, we used a consensus clustering to look for the optimal subdivision of our data into groups as well as for the robustness of each group (see Section 5.6.4). This allowed us to find the groups that better fit the collected data as well as a much more stable solution (Monti et al., 2003).

In Fig. 5.5a we present the results of the clustering of the participants of the control group, with equal capital distribution (equal treatment). In this scenario the optimal number of groups identified is two. Looking at the results of the clustering we immediately identify two different types of behaviors, a group of generous participants (cluster 1) that contribute above the fair contribution, and a group of more greedy participants that contribute around the fair contribution at the beginning of the game and decrease their contribution as they approach the end of the game (but before reaching the goal).

The results for the unequal treatment show that the optimal division of the participants is into three groups, being clusters 2 and 3 those gathering the majority of subjects (92.55%), similarly to the equal treatment. Nevertheless, in this case emerge a minority group (7.45%) of hyper-generous individuals (cluster 1) that contributes far beyond what we are considering fair. And, again, in cluster 2 the subjects contribute on average more than the fair contribution, whereas in cluster 3 the average contributions are around the fair value at the beginning, but as the game approaches the end they decrease below the fair amount.

In this latter framework, participants have different initial endowments. Thus, it is interesting to check how these subjects are distributed in each of the three groups based on their relative

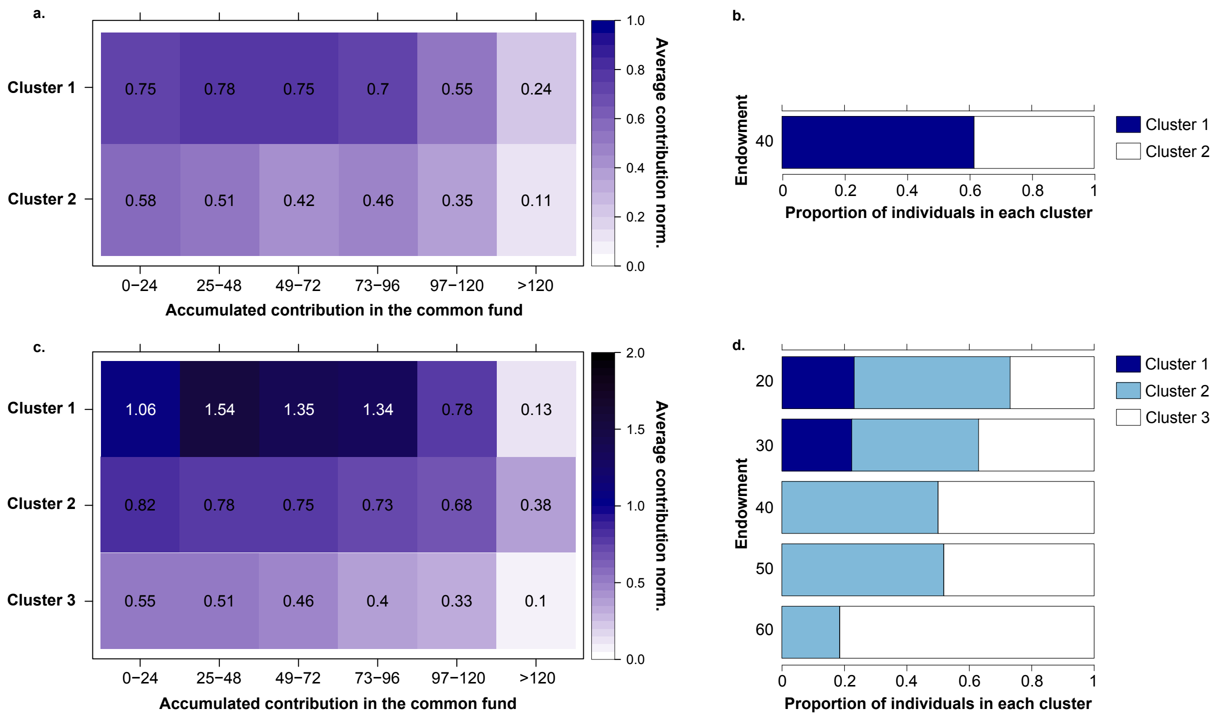

Figure 5.5: a. Behavioral patterns in the equal treatment based on average contribution during the evolution of the game. The value in each cell represents the average contribution normalized by the initial capital per round (i.e: 2 for participants starting with 20 €, 3 for participants starting with 30 €, and so on; 0.5 is the fair contribution) in a given stage of the game (depending on the accumulated contribution at that stage). **b. Proportion of individuals in the equal treatment groups**. Cluster 1 is formed by generous subjects (61%) with average contribution above the fair while cluster 2 is formed by subjects (39%) that contribute around and below the fair contribution. Unequal treatment. **c. Behavioral patterns in the unequal treatment based on average contribution during the evolution of the game**. The value in each cell represents the average contribution normalized by the initial capital per round in a given stage of the game (depending on the accumulated contribution at that stage). Cluster 1 consists of hyper-generous individuals (7.45%) that contribute very much above fair, cluster 2 is formed by generous individuals (43.48%) with average contribution above fair, and cluster 3 is formed by individuals (49.07%) that contribute around and below the fair contribution. **d. Proportion of individuals in the unequal treatment groups**. Distribution of the different types among the participants as a function of their initial endowment.

contribution. Figure 5.5d shows that subjects with fewer resources (20-30 €) than the rest are concentrated in the generous clusters (1 and 2). In fact, the cluster 1 (the hyper-generous group) is formed exclusively by subjects with low endowments. On the other hand, the third cluster is mainly composed by subjects with mid and high endowments. This means that the majority of low endowed participants, 73.07% (20 €) and 62.96% (30 €) contributed above the fair threshold, different from the subjects with high endowments, where only 51.85% (50 €) and 18.51% (60 €) contributed a fair amount. Interestingly, the comparison of Figs. 5.5b and 5.5d shows that subjects with mid endowments (40 - 50 €) distributed among the two clusters not very differently from the equal treatment. Therefore, the richest participants were those who diverged from that distribution.

5.3.4 Effect of awareness about climate change

All groups reached the goal. Although climate awareness has proven to have an effect on individual responses (Burton-Chellew et al., 2013), the majority of our sample (N=294, 91.3%) failed basic questions about climate change concepts, in a questionnaire that included basic questions about the greenhouse effect, carbon footprint, or the Kyoto Protocol (see Appendix B.4). We stress that this is so even if the experiment was done the week following the COP21 summit in Paris, which had led us to expect much more familiarity of our subjects with climate change. This result allows us to exclude that more literacy and more public outreach efforts on climate change are the reasons for the participants reaching the goal in all cases.

5.3.5 Effect of socio-demographics and beliefs

No significant differences were observed in terms of age or gender of the participants. However, educational level was a factor that significantly affected the average contribution (GLMM, χ^2=3.811, df=1, p=.006). Subjects with lower education level were predicted to make higher contributions in equal conditions (F(1,156)=7.219, p<.05). In addition, a third part of the sample (N=112, 34.8%) expected to arrive to the common goal before starting the game. Harboring this previous expectation did not have an influence over their average contribution (χ^2= 6.005, df=3, p=.111).

Importantly, we detect some socially undesirable inconsistencies between belief and behavior. The majority of the sample (around 87%) claimed that their contributions should not depend on the co-participant's contribution, but the existence of different clusters shows that they did took the accumulated capital into account to decide their next contribution. In addition, more than half of the participants (N=214, 66%) defended the idea of relative fairness (e.g. agreed with the statement "Contributions should be proportional to the initial capital so that players with more capital should contribute more to the pool"). In contrast, those who firmly agreed with proportionality in the unequal conditions contributed less than those who did not strongly support that (Z= -2.653, $p < .05$), especially when their initial endowment was high. For example, the participants with a starting capital of 60 € that adhered to "proportionate contributions" contributed an average of 2.6 € per round, whereas those with the same initial endowment who did not firmly claim that contributed 3.05 € per round.

5.3.6 Effect of generosity on emergence of inequality

Finally, we measured what the impact of this highly sacrificed behaviour among the poorest participants on the global inequality was. Gini Coeficcient, apart from a standard way to calculate inequality levels in economics, is completely appropriate here since general behaviour consisting on contributing the fair share would leave the value of Gini Coefficient constant. Thus, in the equal treatment the Gini Coefficient increases from 0 (initial distribution) to 0.1806 at the end of the experiment, whereas in the unequal treatment increases from 0.1812 to 0.3483.

5.4 Discussion

Our experimental results allow us to conclude that inequalities may lead to unexpected problems in climate change mitigation, mostly related to environmental justice. Interestingly, knowledge about climate change facts did not play a relevant role in the average contributions to the common fund. This might indicate that policies mainly addressed to increase climate change awareness might not be the most efficient solution to foster cooperation, and suggests that emphasizing a correct interpretation of the perceived effects might be more useful in this regard.

Even if all groups avoided the climatic catastrophe, our work shows that other potentially serious issues may arise in the process. A particularly important one is that disadvantaged individuals are contributing much more than a fair share of the mitigation, and that the richest ones are contributing less. It is telling that all hyper-generous behavior is observed in the two poorest types of individuals, while a large majority of those endowed with the largest amount behaved selfishly (irrespective of what they claim to believe about fair contributions, as we have seen). It thus appears that, contrary to the expectations of the poor exploiting the rich in a public goods context, here we found the opposite situation. At this point, it is important to note that these experiments were done in a short period of time and therefore they do not inform about long-term behavior. It would be possible that the fact that the poor contributed more for a long time might eventually lead them to stop doing so, thus jeopardizing the collective goal. The decisions we observed in the experiment did not allow us to learn about the behavior of the richest players in that case, i.e., whether or not they would jump in to solve the problem (although the results in Milinski et al. (2011) suggest that this could be the case).

Finally, it is important to discuss the different behaviors observed in the collectives we have worked with. Education does not help here: less educated-less favored participants contributed more to the collective goal than more educated-more favored ones. This could indicate that there is an underlying cultural assumption of sacrifice of the most disadvantaged people (related to their vulnerability): in a situation where the poorest are the ones who will face the worst consequences, more advantaged participants may feel inclined to contribute less to solving the problem. Particularly alarming is the fact that, in the group of the richest participants, about 80% behaved in a selfish manner. As this is the group that had the largest means to help to mitigate climate change, their fault to contribute may jeopardize the whole society, which calls for specific actions to work with this segment of the population while providing additional policies to protect more disadvantaged groups or collectives.

To suggest new ways to initiate a collective climate action according to our findings, firstly it is worth noting some differences of our study with a real world situation. For example, similar to (Burton-Chellew et al., 2013), our participants gathered information about their co-participants' responses whereas no particular information from others (i.e. foreign countries or other communities sharing space in a city) may not be available, reliable, or complete in the real world. Therefore, to foster cooperation, future policies may benefit from making data and contributions public and transparent. Moreover, our experimental subjects held unambiguous responsibility over their actions whereas climate change is a global problem with diffuse shared duties. In this sense we have proven that a good general education is not the remedy to avoid inequality in contribution, but promoting collective rather than parochial attitudes, which has proven to be one factor underlying cooperation (Buchan et al., 2009), may be a better solution to make individuals assume their responsibilities. Finally, we used an economic paradigm to establish a concrete threshold to be reached so that our subjects received the economic consequences directly and immediately once the game was over, however the Nature's threshold is more uncertain (Barrett and Dannenberg, 2012) and consequences can spread out over generations. Monetizing Nature by establishing concrete thresholds to be reached in a particular time period and rewarding the population if they enact some actions (i.e. return taxes if substituting diesel oil cars with electric vehicles or if groups of neighbors install and mantain a rooftop garden on their own buildings) may work to address people towards a general cooperation.

5.5 Methods

The experiment was conducted following the lab-in-the-field experiment guidelines used in (Poncela-Casasnovas et al., 2016), which helped us recruiting participants form a general audience by using opportunistic recruitment, opposed to the typical samples of university undergraduate students. All participants in the experiment signed an informed consent to participate and no association was ever made between their real names and the results, in agreement with the Spanish Law for Personal Data Protection. This procedure was approved by the Ethics Committee of Universidad Carlos III de Madrid. The experiment was conducted in different sessions in the DAU fair in Barcelona during two days (December 12-13, 2015). The total number of games was 54, the number of participants in our experiment was 324 subjects, adding up to a total of 3240 game decisions collected. If some participant was non-responsive, the experimental platform took over and make the contribution randomly for her; in that case the data was labeled and the subject's decisions were discarded in the analysis. The total number of subjects with uncorrupted decisions are 320 (134 women, 41.88%) leading to a total number of valid decisions of 3200. The age ranged from 11 to 73 years, 32.15 (13.04) on average (SD) , there are no significant differences between the behaviour of minors (< 18) and adults at the level of individual contribution, nor between games in which there are minors and adults coexisting and in which there are only adults (see Fig. 5.6). Almost half of the sample was graduated (48%), whereas the other half was equally distributed between different educational levels (i.e. professional training (16%), elementary (11%), middle (11%) and high school (12%). Most of the participants were naïve to social experiments (N=279, 86%). Average (SD) earnings were 18.21€ (8.8) (see Fig. 5.16) and the average (SD) duration of a game (considering only the time of decision making) was 82.21s (20.98), hence the average round time was 8.22s (see Fig. 5.17).

Before starting the game, all the participants were shown a brief tutorial in the tablet in which the experiment was implemented (see Appendix B). Researchers present in the experiment reviewed the instructions with them to guarantee they were understanding the basics of the experiment. Participants were reminded that they had to make a decision on each round on how much money they want to contribute to the common goal, but they were not instructed in any particular way nor with any particular goal in mind.

The subjects participated in groups of six players, each subject was assigned with an initial capital (20 to 60 €), and the goal of the game was to contribute 120€ on a common fund between all of them. The subjects contributed into the common fund (0 to 4 €) during 10 rounds, at the end of each round, all players saw the information of how much money has been contributed to the common fund, they also saw the individual contributions of the six players in the previous round and the total amount contributed by each player up to this round (see Appendix B). If the goal was reached at the end of the game, all the participants kept the money that they had not contributed in the form of a gift card. The 120 € collected in the pot will be used once the research is published in organising an event that includes: (1) an action against climate change (the action will consist in planting trees in a forest in Barcelona within the Day of the Tree event organised by an NGO (NGO, 2016)) and (2) the dissemination of the current results. Otherwise, if the common fund did not reach 120 € at the end of the game, we did not take any action against climate change and the participants only kept the remaining money with a probability of 10%.

5.5.1 Statistical analysis

Numerical data representing payoff and contributions are expressed as mean±SE, in Fig. 5.4 and Fig. 5.3 with a 95% CI. The sociodemographic (age and effect of minors), as well as earnings and time of decision making are expressed as average (SD) and in the respective figures as stardard error of the mean (95% CI). To control for potential sociodemographic data that could have an influence over the participant's responses we used GLMM Baayen (2008) with sex, age, education and the 2- and 3-way interaction between sex, age, education, followed by linear regression post test to establish the direction of the causality found at the educational level. Comparisons between average contribution and qualitative responses to question 3 (reaching the threshold, see Appendix B.4) were conducted with non-parametric Chi Square. To explore how attitudes towards proportionality (question 7, see Appendix B.4) affected contributions, we sorted contributions according to the binomial response in question 7 (group 1 against group 2) and used non-parametric test for independent samples (Wilcoxon-Mann-Whitney test) for comparison. To measure the differences among minors and adults in terms of contributions (both group and individual), we conducted Welch two sample t-test. Finally, to measure the robustness and stability of a given cluster we use consensus cluster metrics as item-consensus and cluster-consensus (Monti et al., 2003).

5.6 Supplementary Information

5.6.1 Sociodemographics

The average age among our valid participants was 32.15 (SD=13.04), with a proportion of 41.88% females and 58.12% males (see Fig. 5.6).

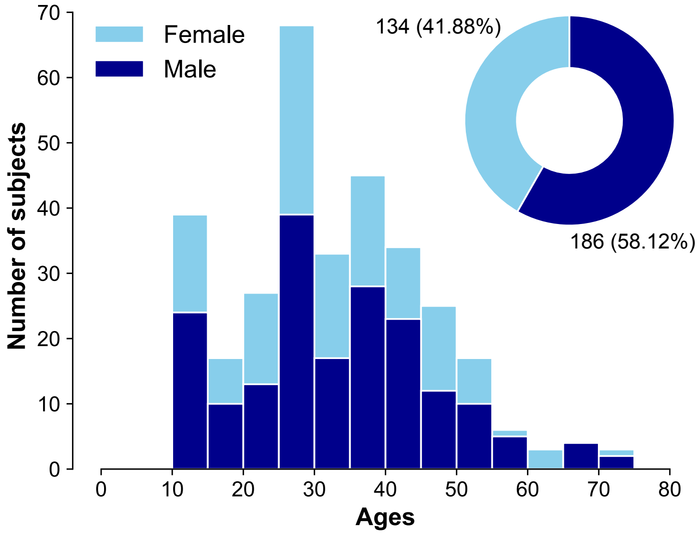

Figure 5.6: Sociodemographics. Distribution of participants in the experiment by age and gender.

Effect of gender in contribution

The average (SD) contribution by gender is: 2.24€ (0.68) in females and 2.13€ (0.72) in males. There are no significant differences in the means by gender (t-test, t: -1.35 p: 0.18).

Effect of minors in games

We consider minors the participants with an age younger than 18. The differences between the behaviour of participants minors and adults, and their influence in the games is described with two parameters: (1) The amount contributed in the common fund in games with minors and adults (29) and games with only adults (25) is 129.7€ (10.2) and 132.2€ (13.4) respectively, no significant differences of the means (t-test, t: 0.8 p: 0.45) (see Appendix Table C.3); (2) in the same line, the individual contribution do not show significant differences for any of the endowment (see Appendix Table C.4).

5.6.2 Selection strategy

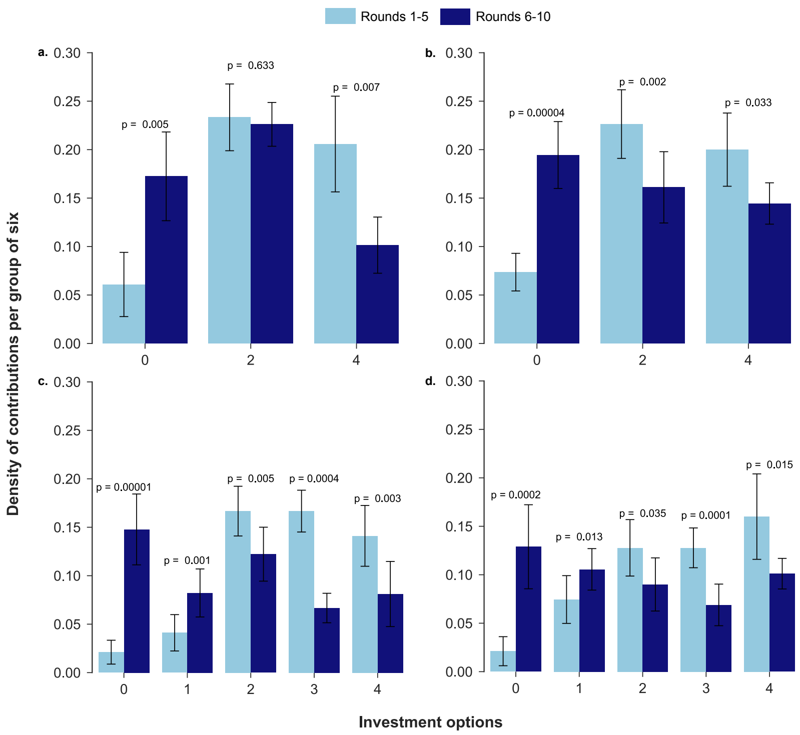

Figure 5.7: Investment choices at the beginning and end of the game. Density of investment selections, mean and standard error of the mean (95% CI), in the first five rounds and the last five rounds. **a.** Equal treatment and investment options of 0-2-4. **b.** Unequal treatment and investment options of 0-2-4. **c.** Equal treatment and investment options of 0-1-2-3-4. **d.** Unequal treatment and investment options of 0-1-2-3-4.

We studied the effect of including more decision options to the participants, instead of allowing to invest 0, 2 or 4 € (control setting) the participants could select to invest between 0 and 4 € (intervention setting). Offering more options to invest allows the existence of more complex investing strategies, since the participants can tune more accurately how much they want to contribute in each round (see Fig. 5.7).

The capital of participants decreased over the game in a different way depending on their endowments. In the unequal treatment, the capital of participants with low endowments decreased faster than the wealthy subjects, partly because the effort to contribute is greater for the poor participants and every single contribution reduced substantially their capital in comparison with the others participants.

The proportion of capital saved in both treatments have no significant difference (t-test $p < 0.05$) if we compare participants with same endowments (see Fig. 5.8a). However we observe in the evolution of the average capital grouped by endowments how adding more options creates a significantly larger gap between the average contribution of richer (50-60 €) and poorer participants (20-30 €). It is also interesting to observe that while the gap is created in the first five rounds, it keeps getting larger until the end of the game (see Fig. 5.8b).

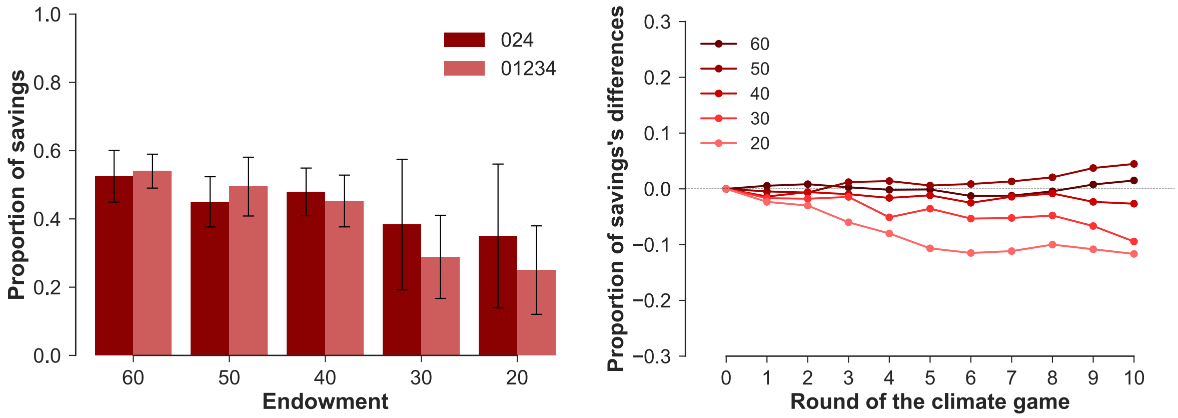

Figure 5.8: Proportion of savings depending on the investment options and the endowments. a. Proportion of savings, mean and standard error of the mean (95% CI), at the end of the game per endowment and investment treatment. **b.** Differences of remaining capital −savings (S)− between the treatment 01234 and 024 per endowment ($S_{01234} - S_{024}$ per endowment in each round).

5.6.3 Game evolution

In our experiment we collected how much one subject contributed in each round. In order to study and infer the existence of behavioral groups or patterns in the data, we first normalized the data so as to find the best classification.

First of all, we normalized the contribution done in each round based on the initial endowments. This is needed in order to compare meaningfully contributions of subjects in the unequal treatment with low endowments (20-30 €), mid endowments (40€), and high endowments (50-60 €), and all of them with the subjects of the equal treatment where all subjects have the same endowment (40€). Note that with this procedure, a normalized contribution of 0.5 represents the fair contribution in all treatments and all endowments.

By looking at the evolution of the game, we observe that each group reaches the target in different rounds −between the 6th and the 10th round (see Fig. 5.9). Therefore, to make the contributions comparable we decided to analyze the evolution of the game not according to the contributions per round, but instead by calculating the average contribution in regard to the accumulated capital in the pot at the beginning of the round.

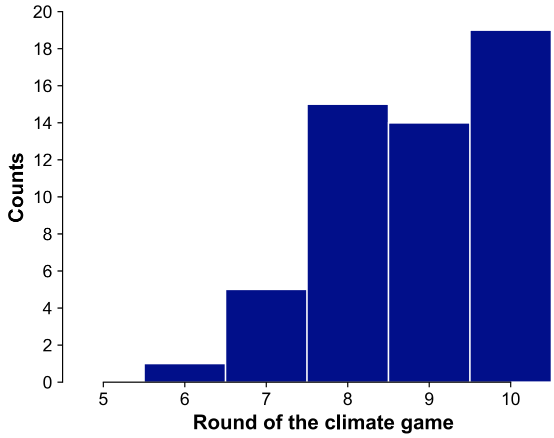

Figure 5.9: Distribution round goal achieved. Number of games in which the goal has been achieved in a particular round. The average (SD) round is 8.83 (1.07).

Phases of the game

The analysis of the evolution of the game according to the amount contributed to the common pot provides a better description of the behavior of the groups, and allows distinguishing different phases of the game. For instance (see Fig. 5.11), we have observed that at the beginning of the game the subjects contribute a lot and explore the contributions of the other participants. However, when they get closer to the goal, there is an important change on how people contribute: their contributions decrease when they see that the goal is close to being achieved (see Fig. 5.10).

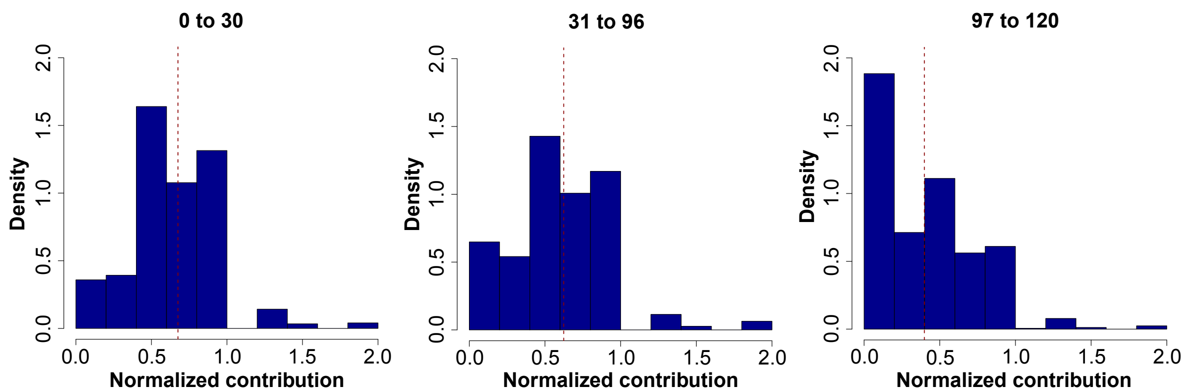

Figure 5.10: Distributions of normalized contributions in the three phases of the game. The mean (SD) in each phase, based on the accumulated capital in the common fund, is: common fund from 0 to 30 €: 0.67 (0.33), common fund from 31 to 96 €: 0.62 (0.37), and common fund from 97 to 120 €: 0.39 (0.38).

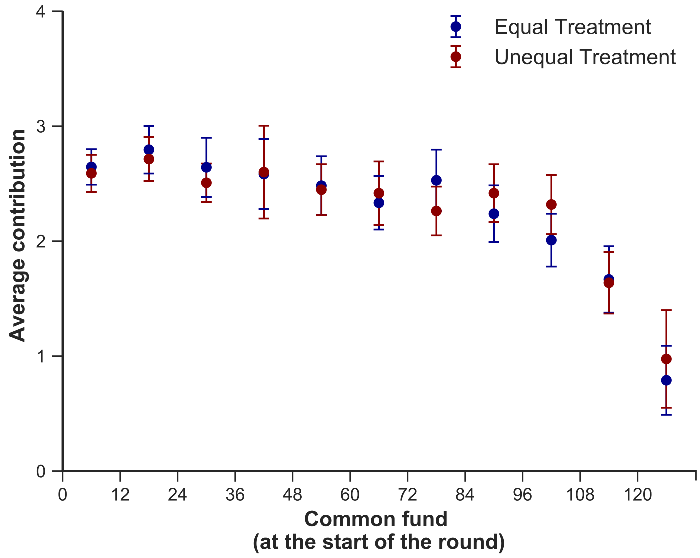

Figure 5.11: Average individual investment and standard error of the mean (95% CI) by treatment over the game evolution (bin=12). Decisions are grouped according to the total capital invested on the common fund at the start of the round. In both equal treatment and unequal treatment participants contribute above the fair contribution in the first part of the game and decrease when they are close to reach the target. We can observe three different regions on the game evolution: first, from 0-30 € participants are more erratic and at the same time contribute more to the average value. Second, from 30 to 90 € approximately there is a stable contribution slightly above the ideal average contribution. And third, after 90 € and until the goal is reached participants decrease substantially their final contribution.

Binning

To study the phases of the game, that is, the evolution independently of the rounds, we bin the rounds according to the accumulated capital in the common fund, and calculate the participants's average contribution per bin and normalize it (see Table 5.1 for an example on how this normalization is done).

5.6.4 Individual Behavior

To study the contribution strategies of the participants of our experiment we first checked whether they follow "pure" strategies by looking at the total of their contributions over the whole game. Note that we only take into account the contributions before the target was reached, and also that the strategies are conditioned by the inequality of endowments. We consider three "pure" strategies: (i) free-riders, those with a contribution of 0€; (ii) fairers, those that have an average contribution per round of 1€ (20€), 1.5€ (30€), 2€ (40€) and so on; and (iii) altruists, those who contribute the maximum that it is possible according to their initial capital, with average round contributions of 2€, 3€, 4€, 4€ and 4€ per endowments of 20€, 30€, 40€, 50€ and 60€ respectively.

We observe that the subjects rarely follow those “pure” strategies, only 42 out of 320 subjects followed a particular strategy: 1 free-rider, 30 fairers and 11 altruists.

Table 5.1: Binning of rounds in Climate Game. Example of user's contribution normalization and binning in a particular game. [1] Contribution of a single user over the game (10-rounds). [2] Capital remaining to achieve the goal (120€ at the beginning of the game). [3] Capital contributed and accumulated in each round. [4] Binning the common fund in groups of 24. [5] Average contribution of a single user in the bin. [6] Average contribution normalized of a single user in the bin.

Round	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Contribution[1]	4	3	4	3	2	2	4	3	3	0
Remaining to target[2]	120	102	82	61	49	26	21	8	1	-5
Common Fund[3]	0	18	38	59	71	84	99	112	119	125
Binning (bin=24)[4]	**0-23**		**24-47**	**48-71**		**72-95**	**96-119**			**⩾ 120**
Av. Cont.[5]	3.5		4	2.5		2	3.3			0
Av. Cont. Norm.[6]	0.875		1	0.625		0.5	0.83			0

In light of the above results and the difficulties to find clear strategies in the experiment, we next ran an unsupervised algorithm to understand the variety of subject's behavior over the game.

Clustering Analysis

We hypothesized that there exist different strategies of cooperation in our dataset that can not be described in terms of pure strategies due to their complexity, but that they could be revealed using unsupervised learning techniques. Hence, we ran hierarchical cluster algorithm on our data to analyze the structure of subject's contributions. We represent the sequence of contribution as explained above (see Table 5.1), creating a matrix of contribution, where every row represents a vector of individual contributions over the game. We use the accumulated capital in the common fund instead of rounds because every game could finish in a different round (see Fig. 5.9), and a bin = 24 due to the rapid evolution of the game. Therefore, in each cell, we have the average contribution normalized (regarding the endowment) in the round that grouped every bin.

We implemented an agglomerative hierarchical clustering strategy (using the hclust package in R) to find an initial approximation of groups using ward.D2 and euclidean distances. The agglomerative strategies consist of a "bottom-up" approach, each user contribution starts in its own cluster and pairs of clusters are merged in each step based on the optimal value of an error sum of squares (Ward's method). With the objective to determine the number of groups that better fits with our data, we ran an algorithm (NBCluster package in R) that computes 26 indexes and recommends the optimal number of clusters according to the majority rule.

Once we know that there is sufficient evidence to find groups of participants with different strategies and behavior, and to ensure that the clustering results are robust and reliable we ran an implementation of consensus clustering (ConsensusClusteringPlus package in R). Consensus clustering determines the number of clusters and computes consensus values such as item consensus and cluster consensus. Item consensus represents the membership of an item with all items in a particular cluster, this value indicates if a particular item is a pure member of the cluster or if it is unstable. Cluster consensus provides information about the consensus between members of a group, high values indicate high stability.

The parameters we used to perform the calculation illustrated in the next sections are: maximum evaluated k of 9 so that groups count of 2 to 9 are evaluated; 1000 re-samplings, agglomerative hierarchical clustering algorithm, euclidean distances and ward.D2 linkage.

Equal Treatment

The equal treatment of the experiment includes 162 subjects (27 games), of which 159 subjects contributed with valid actions. In order to analyze their individual strategies we created a matrix with the average contribution in each stage of the game binned by the accumulated capital in the common fund, and then we computed the cluster consensus in our equal treatment dataset.

The number of clusters that better fits with our data and maximizes the consensus values is 2 (see Fig. 5.12). The average (SD) consensus clustering ration for 2 groups is 0.75 (0) and the average of item consensus is 0.75 (0.08). Cluster 1 is composed by 97 subjects (61%) while cluster 2 is formed by 62 subjects (39%). Fig. 5.14 represents the distribution of subjects (pdf) and the cumulative distribution function (cdf) of their average contribution in both clusters.

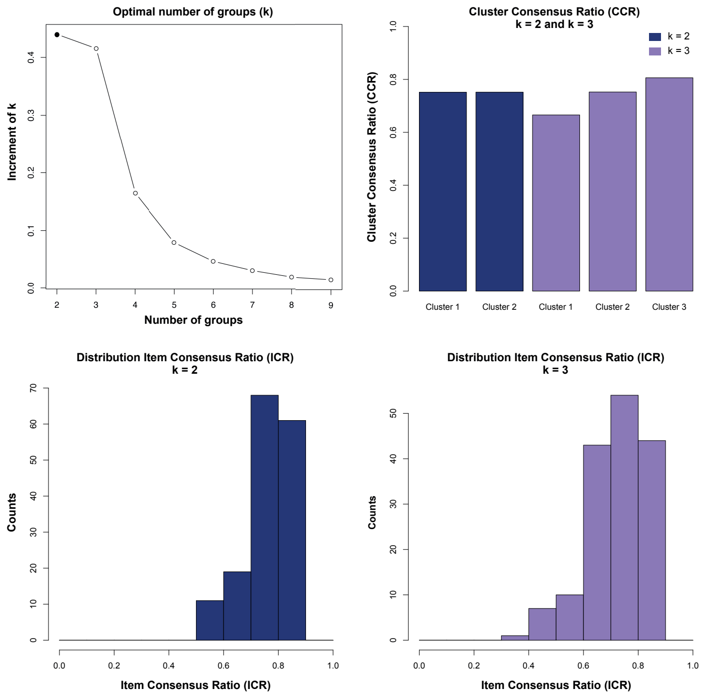

Figure 5.12: Cluster metrics in the equal treatment. a. Optimal number of clusters. **b.** Cluster consensus ratio. **c. and d.** Item consensus ratio.

Unequal Treatment

The unequal treatment includes 162 subjects (27 games), of which 161 contributed valid data. The matrix of contributions was formed in the same way that in the equal case.

The consensus cluster approach concluded that 3 clusters is the most stable number of groups in this treatment, see Fig. 5.13. In this case, the average (SD) consensus clustering ratio is 0.83 (0.12) and the average of item consensus ratio is 0.78 (0.09).

In contrast to the equal treatment, in the unequal scenario a new group appears. The composition of clusters is: cluster 1, 12 subjects (7.45%); cluster 2, 70 subjects (43.48%); and cluster 3, 79 subjects (49.07%). The most populated cluster is the one composed by subjects that contributed around and below the fair contribution. Figure 5.15 represents the distribution of subjects (pdf) and the cumulative distribution function (cdf) of their average contribution in both clusters.

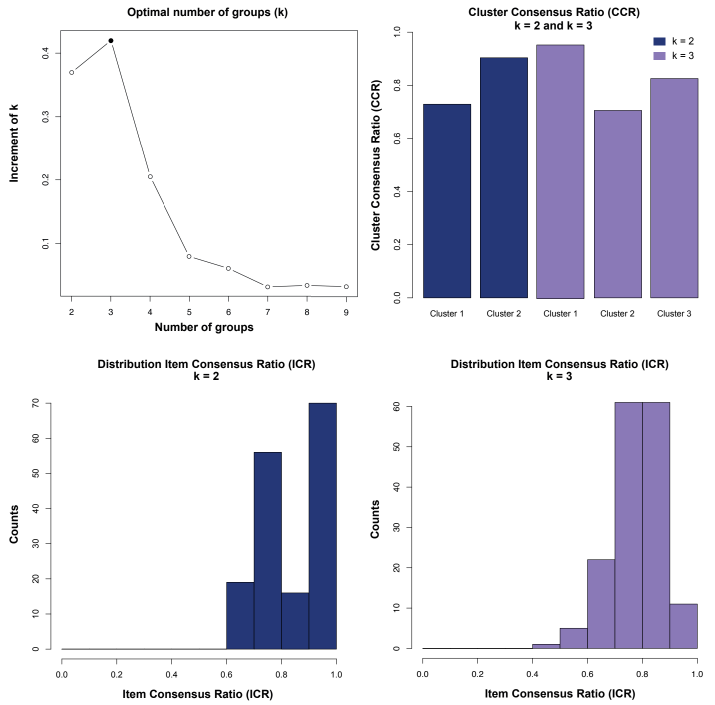

Figure 5.13: Cluster metrics cluster in the unequal treatment. a. Optimal number of clusters. **b.** Cluster consensus ratio. **c and d.** Item consensus ratio.

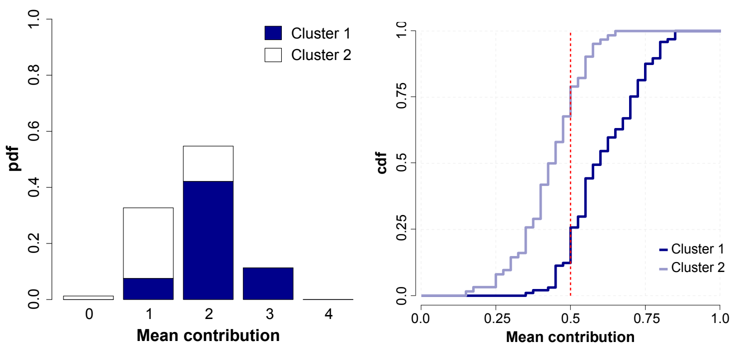

Figure 5.14: Distributions in clusters equal treatment. a. Distribution of subjects in clusters based on their average contribution per round. **b.** Cumulative distribution function based on their average contribution per round.

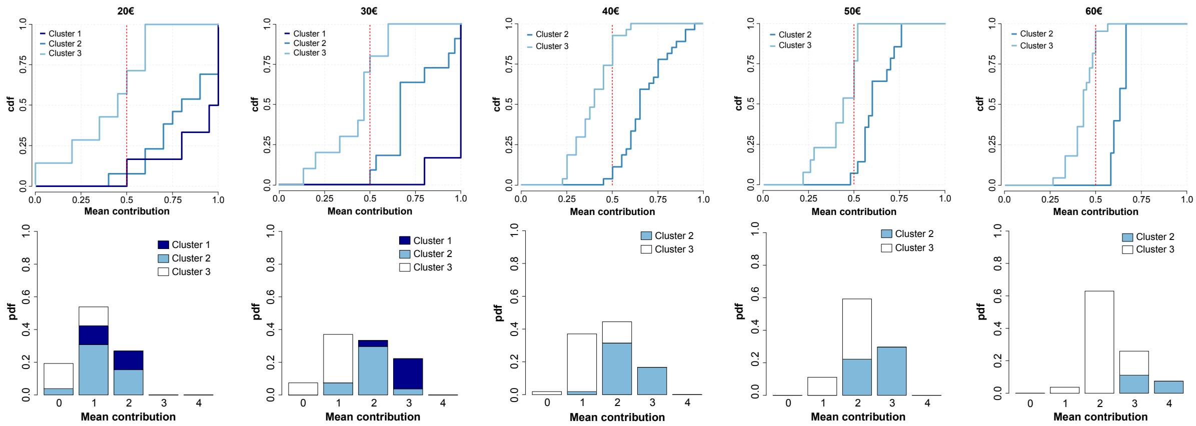

Figure 5.15: Distributions in clusters unequal treatment. (Top) Cumulative distribution function based on their average contribution per round. **(Bottom)** Distribution of subjects in clusters based on their average contribution per round.

5.6.5 Earnings

Once the game finished and if the subjects achieved the goal, they kept the capital not contributed and received it in the form of a gift card. The average (standard deviation) earnings among all the subjects were 18.21€ (8.8). There exist no significant differences in the means (t-test; t:0.26, p:0.8) between the equal treatment 18.33€ (5.8) and the unequal treatment 18.08€ (11). Figure 5.16 illustrates earnings regarding of endowments and treatments in detail.

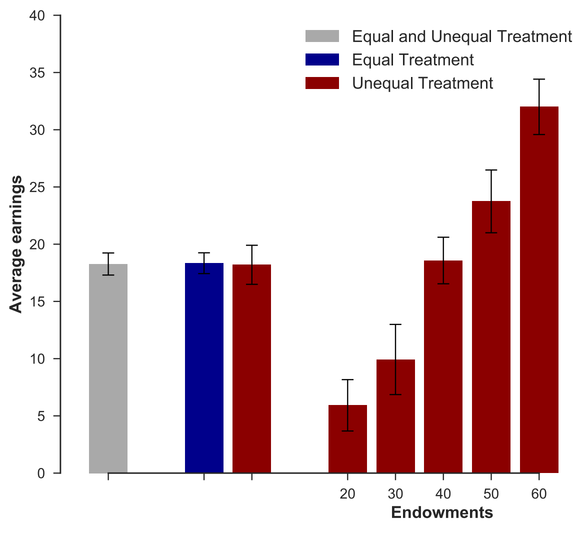

Figure 5.16: Earnings. Average earnings and standard error of the mean (95% CI) regarding treatment and endowments

5.6.6 Decision Making Times

The platform designed to carry out the experiment also records the decision making times and the duration of a round, and we can also obtain the total playing time by summing the durations of rounds in a game. The average (standard deviation) duration of all games was 82.21s (20.98), and the average (standard deviation) duration of a round was 8.22s (4.47).

There are significant differences in the mean (t-test; t:2.41, p:0.019) between equal treatment 88.88s (25.01) and unequal treatment 75.55s (12.89). The decision times decrease as rounds go on, especially in the first five rounds, and subsequently stabilizes (see Fig. 5.17).

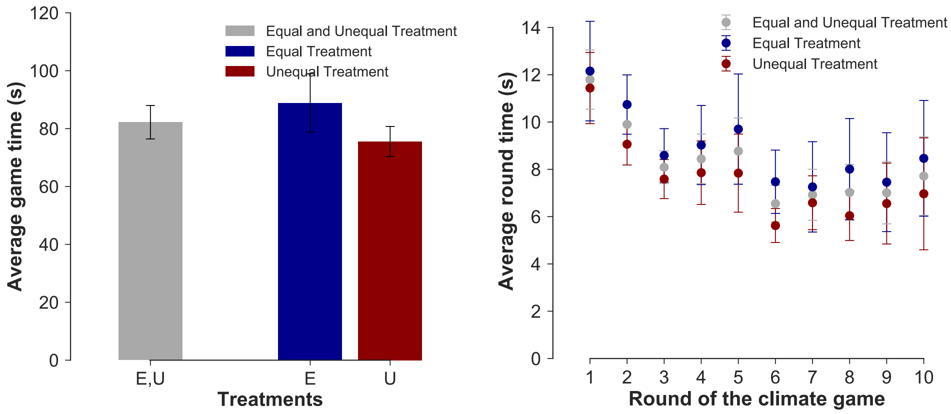

Figure 5.17: Decision making times. a. Duration of a game, mean and standard error of the mean (95% CI), per treatment. **b.** Evolution of decision making times over round.

Chapter 6

Collective Sense in the Mental Health Ecosystem

Summary – Mental disorders have an enormous impact on our society, both personally and structurally. We analyze quantitatively the importance of communities for effective mental health care, considering all community members involved. The experiments has been performed in different locations using the experimental platform presented before. By means of citizen science practices, we have designed a suite of games that allow to probe into different behavioral traits of the role groups of the ecosystem. Here, we present a collective-risk dilemma that analyses "collective sense", a particular behavioural trait. The evidence shows that the cost of collective action is mainly supported by individuals with a mental condition - which unveils their vulnerability. We point to the conditions under which the experiments are carried out in a socialized context and conclude that they can be applied in any similar "care in the community" initiative.

6.1 Introduction

Approximately one fifth of the world population will suffer some mental disorder (MD) at some point in their lives, such as anxiety or depression (Steel et al., 2014). The direct economic costs of MD, including care and indirect effects, is estimated to reach $6 trillion in 2030, which is more than cancer, diabetes, and respiratory diseases combined (Insel et al., 2015). As part of a global effort to scale up services and bring down costs, reliance is increasingly made upon informal social networks (White et al., 2017). A holistic approach to mental health promotion and care provision is then necessary, and emphasis is placed on the idea of individuals-in-community: individuals with MD are defined not just alone but in relationship to others (World Health Organisation, 2001). Such a paradigm shift implies superseding the traditional physician-patient dyad to include caregivers, relatives, social workers, and the community as a whole, recognizing their crucial role in the recovery process.

A key aspect in the definition and aetiology of MD has to do with social behavior (American Psychiatric Association, 2000): behavioral symptoms, or consequences at the behavioral level, characterize most MD. For instance, autism, social phobia, or personality disorders are determined by the presence of impairments in social interaction. Other disorders result in significant difficulties in the social domain, such as depression or psychotic disorders. Further, conditions that are intrinsically behavioral (as for eating disorders or substance abuse) seem to be exacerbated by the influence of social peers. A large body of research has therefore looked at the neural basis of social decision-making among individuals with MD to identify objective biomarkers that may prove useful for its diagnosis, therapy evaluation, and understanding (Gradin et al., 2015; Shao et al., 2015; Guroglu et al., 2010). However, such a methodology does not well fit into the individuals-in-community paradigm. We argue that an agent-based approach which draws upon experimental game theory might prove insightful and ecologically valid for the study of behavior in a given social environment.

Within the mental health literature, the use of game theory as a way to understand the multi-faceted dimensions of behavior has received already quite some attention (Wang et al., 2015; King-Casas and Chiu, 2012). Most research addressed the issue of behavioral differences between individuals with MD and healthy populations (Gradin et al., 2015; Shao et al., 2015; Pulcu et al., 2014; Wang et al., 2014; Csukly et al., 2011; Agay et al., 2008; Mokros et al., 2008; Rilling et al., 2007). These works, that point to cognitive and affective processing impairments (Gradin et al., 2015; Rilling et al., 2007; Scheele et al., 2013), further support the idea that MDs are associated with significant and pervasive difficulties in social cognition and altered decision-making at various levels. Yet, despite these studies are of very much interest, they are primarily concerned with dyadic interactions among people with specific MDs. That is, they lack insights into the complexity of individual behaviors of MD within a specific social context.

Here our objective is to provide quantitative accounts of sense of community in the mental health ecosystems. Thus, we consider the heterogeneity of actors that participate in the mental health ecosystem. We strongly believe that to understand the collective behaviour of a community is necessary to consider the relation context of the individuals. In real social context the interaction among individuals occurs in diverse forms: sometimes people have to coordinate or anti-coordinate their efforts, they find themselves in situations of inequality, yet in other situations they have to work together to achieve a goal. Collective sense, this is the trait in which we focus on, which calls for an experimental approach in which put together actors of the mental health community to interact collectively to solve a particular problem.

For these purposes, we have designed an experimental setup that probes into the complexity of the interdependencies at play within the mental health ecosystem. Accordingly, our experiments take place in a socialized, lab-in-the-field setting (Sagarra et al., 2016), in order to be as close as possible to the dynamic and unique nature of real-life social interactions. We have developed a framework that allows to capture some difficult-to-observe aspects of behavior and social capital within mental health ecosystems as a way to understand how communities contribute to care and resocialization. Here we present part of the framework addressed to understand collective behaviour in the mental health ecosystem.

6.2 The collective risk dilemma

Participants participated in a suite of social dilemmas, here we present, one of them, the Collective Risk Dilemma (Milinski et al., 2008). Briefly, the game is a public goods game with threshold: Each participant owns 40 MUs and contributes in a common found, if the participants' total contribution after 10 rounds is lower than a given threshold (120 MUs), they loose all the money they kept with a probability of 90%. Otherwise, they are told that the money collected in the common fund are spent in reforesting land plots in Catalonia, where the experimental sessions took place, and each participants earns the money left in the personal account. Therefore, the participants played a basic version of the game, in which the whole group had to reach a common goal to avert a catastrophe that most likely would wipe out their money. Participants belonging to the mental health ecosystem played with each other in group of six players. However, they could by no means guess with whom they were actually playing.

6.3 Results

We present the results of the collective-risk social dilemma, one of the suite of games in which the participants contribute with their actions, we focus our analysis in the group interaction among the participants.

6.3.1 Group interaction

Our experimental setup has proven extremely informative in its most novel section, namely the analysis of group interactions framed within the Collective Risk Dilemma (CRD), with no prior result within the mental health literature. In global terms, the average amount contributed to the public good (22.6 MUs) is much more than the fair contribution of 20 MUs, where by fair we understand sharing equally the total amount needed for the threshold (120 MUs) among all six participants. Here it is important to keep in mind that participants were told that all money contributed would go to reforestation projects, so it is not irrational to keep contributing beyond the threshold as many of our subjects did. The key result in the CRD is that large, significant differences (t-test, $t = 2.85, df = 242, p = 0.0047$) are found between participants with and without mental disorders. The former contribute with 22.95 ± 0.63 MUs compared to 20.34 ± 0.68 MUs from the latter, and therefore it appears that when repeated interaction and sustained teamwork (CRD) are required, people with MD contribute much more to the common goal.

Contribution dynamics vary according to group composition in terms of number of participants with mental disorder conditions and other actors involved in the recovery process. All groups successfully reach the target collecting on average 135.64 ± 1.75 MUs. Similarly to other public good experiments, contributions decrease over time (Milinski et al., 2008). While in the first round participants contribute around 56.3% of the allowed contribution per round (2.2 ± 0.07 MUs, where the social optimum is 2 MU), contributions drop when the endgame effect sets in. A Spearman's rank-order correlation of contributions over rounds corroborates this negative time trend ($\rho = -0.757, p < 0.05$). Both patients and actors involved in the recovery process reduce their contributions by the end of the game. However, in almost all rounds, participants with a mental condition contribute more than caregivers and non caregivers, for whom motivations to contribute decline steadily (see Fig. 6.1).

In terms of the group composition, groups where individuals with MD conditions constitute half or the majority of the group (n=36) do much better in sustaining cooperation compared to groups where firsthand affected are the minority (n=9). It is here worth to mention that participants may see who the rest of the members are but ignore who is exactly making the choice in the game (see Methods for further details). As Fig.6.2b shows, while average individual contributions are similar in the last periods (rounds 6-10 t-test, $t = 0.19$, $p = 0.85$), groups with half or more individuals with MD contribute significantly more at the beginning of the game (rounds 1-5 t-test, $t = 2.79$, $p = 0.0054$). Hence, the presence of three or more individuals with a mental condition in the group has a positive and stabilizing effect on average individual contributions. Likewise, in games with a low proportion of participants affected with MD the group achieved the goal, on average, later than in games with more than 50% of participants affected with MD (see Fig.6.2a).

If we then break down the analysis by group type, we find that group members contribute and benefit differently from cooperation (see Fig.6.2c). Indeed, final payoffs within groups are far from being equally distributed (see Fig.6.2d), with the highest inequality found in the group where the number of patients equals the number of actors involved in the recovery process (Gini

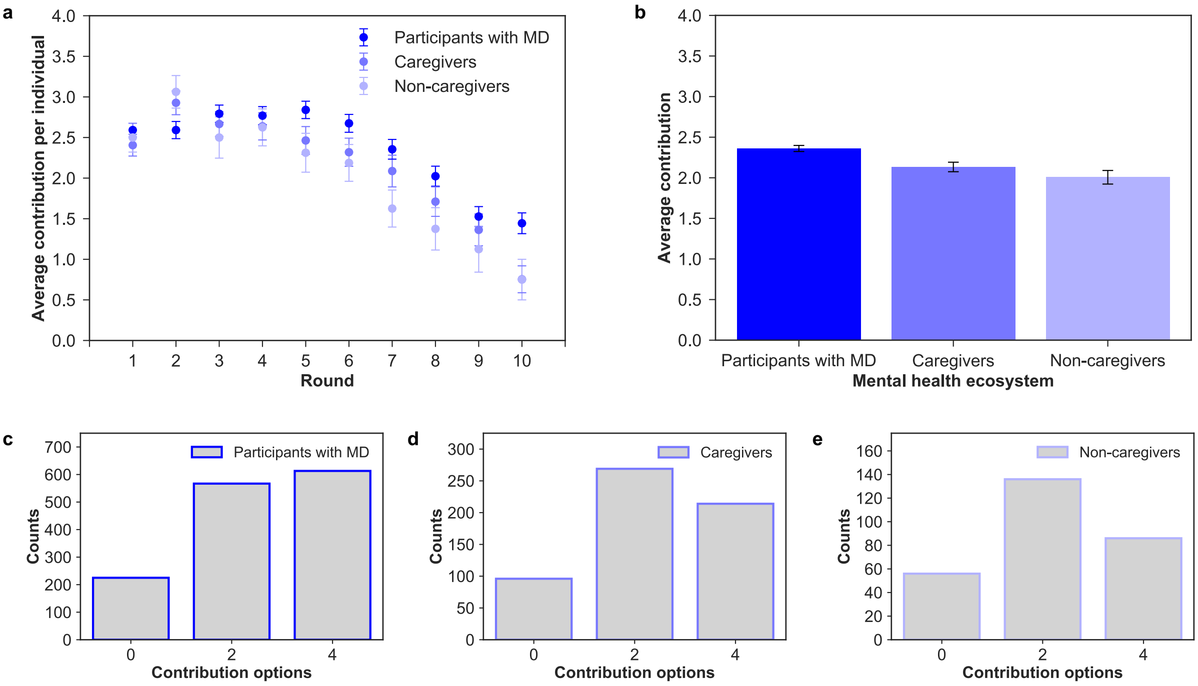

Figure 6.1: a. Individual contribution over rounds. Evolution of contributions (mean and standard error of the mean) during the game between participants with mental disorder conditions, caregivers and non-caregivers. We can see that all groups behave similarly and in an identical way to a previous experiment run outside the mental health ecosystem (Vicens et al., 2017). **b. Average individual contribution per round.** Average contribution and standard error of the mean in the mental health ecosystem. There are significant differences between participant with MD and the rest of actors, caregivers (t-test, $t = 2.107, df = 155, p < 0.0294$) and non-caregivers (t-test, $t = 2.499, df = 48, p = 0.01588$). **Distribution of choices by: c. participants with MD, d. caregivers and, e. non-caregivers.** The most of participants with MD (43.6%) selected the maximum contribution (4), while the caregivers (46.5%) and non-caregivers (48.9%) mostly selected the fair contribution (2).

coefficient = 0.289). We thus see clearly that the cost of collective action is mainly supported by individuals with a mental disorder. Given that they contribute the most within all groups, lower investments are needed for other members of the collective to reach the common target. Yet, in 4/6 and 5/6 groups caregivers reduce average individual contributions while non-caregivers pay more than their fair share. In 1/6 and 2/6 groups, on the other hand, caregivers are the ones who compensate the unfair contributions of other members. These last groups are the ones that ensure the lowest inequality in final payoffs. Therefore, while our results are unambiguous about the larger readiness for collective action among people with MD, we cannot claim nothing about the rest of the collective.

6.4 Discussion

As a first general remark, through our lab-in-the field experiment we found that an ecosystem approach to mental health care brings with it a quite complex scenario with several interesting insights. To begin with, members of the mental health ecosystem do not equally contribute and

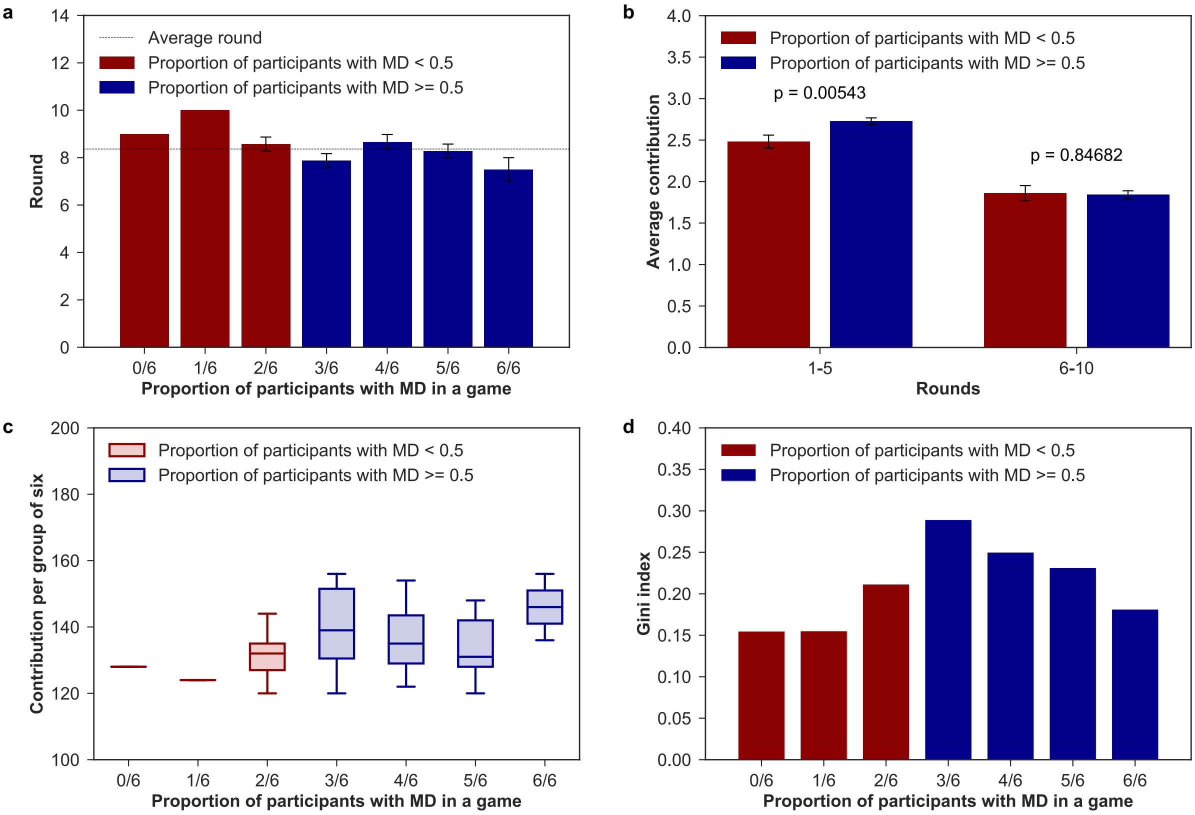

Figure 6.2: **a. Average round of achievement.** Round (mean and standard error of the mean) in which the group of six achieved the target. **b. Aggregated contributions per group composition.** Contributions (mean and standard error of the mean) in the first and last five-rounds per number of individuals with MD in a group. There are significant differences (t-test $p < 0.01$) in contributions in the first part of the game. **c. Contributions per group of six.** Total group contributions by number of individuals with mental conditions in the group. **d. Gini index of final payoff within groups.** Level of inequality in final payoff based on the number of individuals with MD in each group.

benefit from collective action. Rather, systematic behavioral differences arise as the number of social interactions increase, i.e., when teamwork is required for the collective to benefit as a whole. This suggests that considering repeated games may prove extremely insightful for the purpose of the research. Indeed, our experiments show that individuals with MD are the ones who contribute the most to the public good: they make larger efforts towards reaching the collective goal, thus playing a leading role for the functioning of the ecosystem. As a consequence, groups with half or more participants with MD do better in sustaining cooperation in the first rounds, which implies that a community care setting might prove successful for capability building. Yet, large proportions of individuals with MD in a group result in higher inequalities in final gains, which reach the maximum when the number of individuals with MD equals the number of caregivers or relatives. This means that community care perspectives might also take account of group composition to deal with potential inequalities arising from differential capabilities.

The larger readiness of individuals with MD to contribute to the collective action problem can thus be seen as a way to claim their place in the community. By having participants unaware of their partner's identity, we could indeed measure participants decisions based solely on the value they placed on the group's welfare, independently of its composition or other factors. Yet, the fact that participants with MD contribute the most implies for other members of the group lower investments to reach the common target. This, on the other hand, unveils the vulnerability of individuals with a diagnosis of MD. Repeated or periodic and more situated experiments with digital platforms (Aledavood et al., 2017), in the future, can surely provide further valuable insights into the effect of participants prior knowledge of and relation with the partner on their behavior.

Here we present the results and discussion of a part of the framework with a suite of games to study the behaviour in mental health ecosystem. We are indeed sure that our complete experimental setup can prove helpful in complementing the diagnostic process of physicians and health professionals and even to evaluate care service providers. On the other hand, other possible application of this approach arises in the realm of behavior change interventions (Blaga et al., 2017), that should focus on the aspects that are more specific of every disorder.

In summary, we have explored the behavior of all individuals and role groups who make up the mental health ecosystem through an extensive suite of games that simulate strategic social situations, and here we present the impact in one particular behavioural trait: collective sense (or collectivity). Indeed, the behavior of individuals with MD can be better explained by examining not only their cognitive abilities, but also the web of relationships in which they are embedded. Yet, that web of relationships presents opportunities and imposes constraints. Finally, given that our work has been carried out in a fully socialized context, this approach can be applied to any similar 'care in the community' initiative.

6.5 Methods

All participants were fully informed about the purpose, methods and intended uses of the research. No participant could approach any experimental station without having signed a written informed consent. The use of pseudonyms ensured the anonymity of participants identity, in agreement with the Spanish Law for Personal Data Protection. No association was ever made between the participants real names and the results. The whole procedure was approved by the Ethics Committee of Universitat de Barcelona. All methods were performed in accordance with the relevant guidelines and regulations.

6.5.1 Experimental design

The dialogue with the main stakeholders of the mental health ecosystem was at the centre of the project. Around 20 representatives including members of the Catalonia Federation of Mental Health (Federació Salut Mental Catalunya), firsthand affected, relatives, caregivers, and other professionals related to both the health and social sector, informed and validated the whole research through focus groups and further discussions, leading to the largest experiment of this kind ever carried out. Citizen science principles guided the whole experimental design process in order to raise concerns grounded in the daily life of mental health professionals and service users, and to increase public awareness. The locations where the experiments took place were accorded with the Catalonia Federation of Mental Health in an attempt to explore the functioning of some communities of interest for inclusive and effective policy making. The Federation provided a fundamental support throughout the whole experiments' implementation, serving as a crucial intermediaries between the scientists and different mental health collectives. It also provided valuable insights to better interpret the data obtained.

6.5.2 Participants and procedure

To our knowledge, experimental work on this issue has been conducted only recently and on specific collectives of orders of magnitude smaller. A total of 270 individuals participated in the experiments, that were run over 45 sessions between October 2016 and March 2017. The experiments were carried out in Girona (n = 60), Lleida (n = 120), Sabadell (n = 48) and Valls (n = 42). Participants were either diagnosed with a mental condition (n=169) or members of the broader mental health ecosystem (n=101), including professionals of the health and social sector (n=52), formal and informal caregivers (n=17), relatives (n=9), friends (n=4) and other members of the collective (n=19). Individuals with a mental condition had to self-assess their diagnosis selecting one from a spectrum of options agreed upon with representatives of the mental health ecosystem during the co-design phases of the experiment. Those participants who had received more than one diagnosis had to select the one they considered to be the most relevant. Overall, they had received a diagnosis of psychosis (n=63), depression (n=33), anxiety (n=31), bipolar disorder (n=17) or other unspecified diagnosis (n=25). They ranged in age from 21 to 77 years old (these are weighted values since for ethical and privacy reasons participants were only asked to choose among different age ranges) with 47.2 years on average. Further, 55.6% were men and 44.4% were women. Yet, actors involved in the recovery process were predominantly women (76.2%), and up to 21.8% of them was over 60 years old. Participants were told that they would play against each others a set of games meant to explore human decision-making processes. They played in random groups of six players through a web interface specifically developed for the research. They were informed that they had to make a decision under different conditions and against different opponents in every round. Every game represented an interactive situation requiring the participants to make a decision, the result of which depended also on the opponent's behavior. To incentivize the participation, they would earn a voucher worth their final score (the experimental settings and instructions, can be found in the Appendix D). First, participants participated in a Collective Risk Dilemma (Milinski et al., 2008) against five opponents. After completing the task, participants played the other games (Trust Game (Berg et al., 1995) and Prisoner Dilemma (Rapoport and Chammah, 1965)). Before starting the games, participants had to complete a brief survey covering some key dimensions of their sociodemographic background.

The average (standard error of the mean) time for completing the collective-risk dilemma is around 4 minutes, $250.14 \pm 4.59s$, and the whole experiment (CRD, PD, TG and tutorials) is $705.86 \pm 17.93s$. The average individual earnings in the collective-risk dilemma is 18.03 ± 0.47 MUs. At the end of each session, that is after the participants complete the set of experiments, they received a gift card worth their earnings. The average individual earning in the complete experiment is 46.84 ± 0.77 MUs equivalent to a 4.04 ± 0.077 € voucher.

6.5.3 Statistical analysis

Results were analyzed at two levels: first, we tested for behavioral differences between the whole group of individuals with mental condition compared to members of the mental health ecosystem; we then checked for systematic behavioral variation across diagnostics and role played in the recovery process. We checked for marginal differences within groups using Kruskal-Wallis tests, and post-hoc comparisons were run with Mann-Whitney-U tests adjusting for p-values with the Holm-Bonferroni method. Welch's two-tailed t-tests were performed to check for differences in average contributions (CRD) between participants with and without a MD, controlling for unequal variances and sample sizes. Finally, ANOVA and further Tukey HSD post-hoc comparisons served to check for differences in average contributions over round across diagnostics and members of the mental health community.

6.6 Supplementary Information

6.6.1 Sociodemographics

Figure 6.3 summarizes some key background figures of our 270 participants. It is noteworthy that although anybody aged over 18 could participate in the experiment, only a small fraction (7%) of participants with a mental disorder were under 30 years old. This implicitly suggests that younger individuals with mental disorders may not be well represented within mental health communities. Also, few were working (20%) and most were not in a stable relationship (55%), with more than half with primary or no education (53%). On the other hand, actors involved in the recovery process were more heterogeneously distributed across age groups. Yet, only 28% were either married or in a civil union, and 43% were not in a paid job.

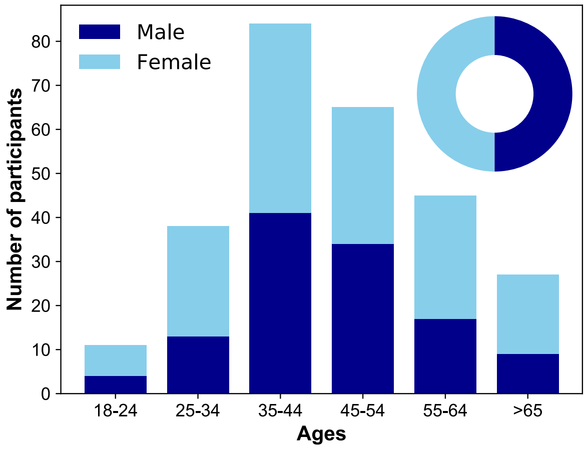

Figure 6.3: Sociodemographics. Age and gender distributions. Age ranges are those given in the survey following Ethics and Privacy committe advice. There were 270 participants: 55.6% were men and 44.4% were women.

6.6.2 Experimental settings

All experimental sessions were carried out in the infrastructures provided by the Catalan Federation of Mental Health. These include: the World Mental Health Day event, organized in Lleida on October 8th 2016 in the background of the Old Cathedral of Lleida. Two social events promoted and arranged by the Federation as part of the yearly social activity plan: in an assembly hall in Valls on March 18th, 2017, and in a community mental health center in Sabadell on March 27th, 2017. And a local employment insertion centre for people with mental disorders in Girona on March 24th, 2017. The participants played in random groups of six players each, for a maximum of three groups at a time depending on the location's constraints. Experimental stations were spatially arranged so that participants could not see each other. Also, they were rigorously prevented from talking or signalling one another. To further guarantee that potential interactions among players sitting close to each other did not influence the results of the experiment, the assignment of players' partners was completely random. All of the participants played through a web interface specifically developed for the experiment. The participants were shown a brief tutorial, but were not given any clue. They were informed that they had to make decisions under different conditions and against different opponents in every round. Also, to incentivize their participation, participants were told that they would earn a voucher worth their final score. We made sure the interface be the most simple and understandable to ensure the correct understanding of the tasks. Also, the interface was the same for everybody. We made sure to avoid the research be upsetting or harmful for the participants by presenting the experiment as a game and playful activity. Three to four researchers closely monitored each session to guarantee the experimental protocol be strictly followed. Yet, the researchers provided help when required.

6.6.3 Collectivity in the Collective-Risk Dilemma

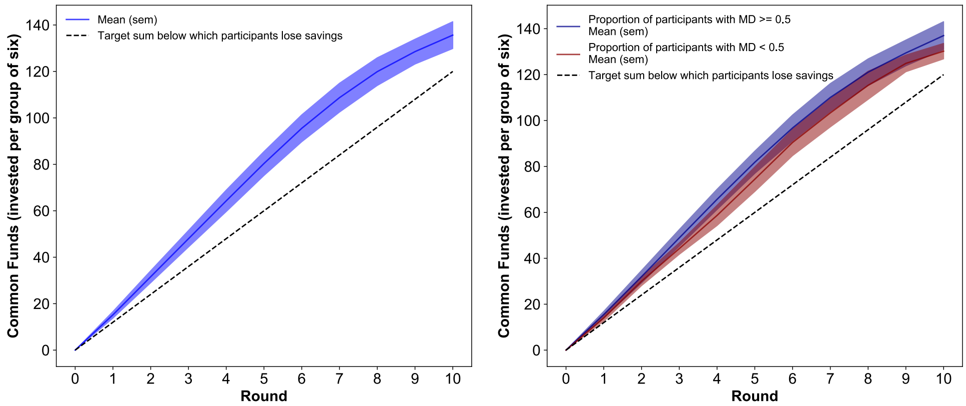

Figure 6.4: Evolution of contributions during the games. (Left) Evolution of aggregate contributions (mean and standard error of the mean) to the common fund over rounds. (Right) Evolution of aggregate contributions (mean and standard error of the mean) over rounds depending on the portion of firsthand affected within groups.

Participants played in group of six players a Collective-Risk Dilemma over 10 rounds. They were endowed with 40 MUs and asked, in each round, to simultaneously contribute to a common fund with 0, 2, or 4 MUs. They had 20 seconds to made their decision after which the computer would make it for them. At the end of each round they would receive information about: the amount of money collected by the group members over the earlier rounds; individual and group members contributions in the previous round; and individual retained MUs. They were told that if the group's total contributions at the end of all rounds reached or surpassed a target amount set at 120 MUs, a certain collective action to mitigate climate change would be promoted and all group members would gain their individual retained funds. If insufficient contributions were made, the contributors would loose their contributions with 90% probability, and the collective action would not be promoted. A social dilemma arises in that everybody benefits from reaching the target sum, but players are tempted to contribute 0 and to benefit personally at the expenses of other group members. The Collective-Risk Dilemma is the paradigm that best captures the social dilemma that emerges from the conflict between group and individual interests. It allows for the realistic modelling of group interactions and individual sacrifice for the group's welfare. Indeed, the more a participant invests in the collective good, the higher the probability that the group reaches the target sum. Yet, the less money remains in his or her personal account. In contrast, failure to reach the target sum implies a high risk that the remaining money in the personal account will be lost. At the same time, the more others invest, the less a subject needs to invest for the group still to reach its target sum. We thus define collectivity as individuals' contribution to the group's welfare. We measure it as the portion of the amount contributed to the common fund over rounds out of their initial endowment (40 MUs).

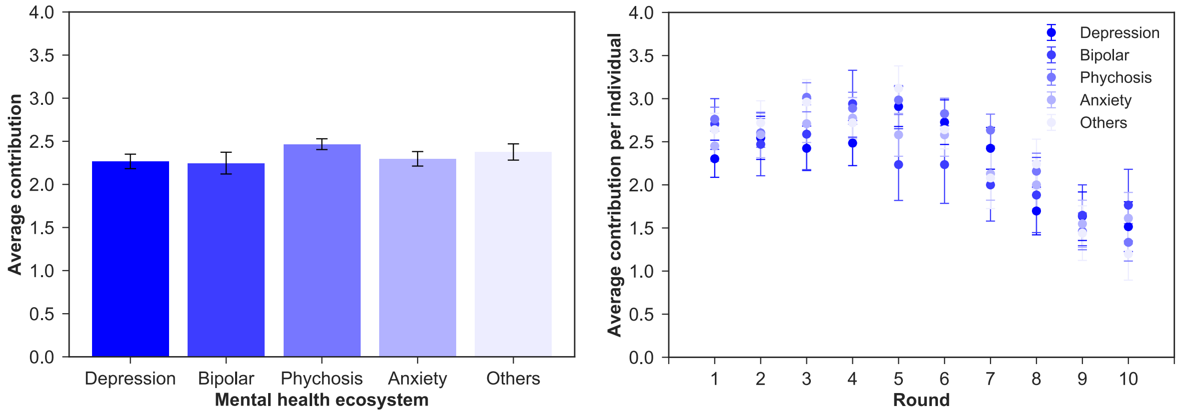

Figure 6.5: Contributions across diagnostics. (Left) Average aggregate contributions (mean and standard error of the mean) and (Right) evolution of contributions (mean and standard error of the mean) over rounds.

Game evolution

Groups of six participants contributed in the common goal during 10-rounds with the objective to collect 120 MUs at the end of the game. As we observe in Fig.6.4, the average contribution of the groups of six in each round is above the fair contribution during all the game, therefore in average the groups always are in disposition to achieve the target. We can observe how in groups with different proportion of participants with MD the trend doesn't change.

Effects of relation with firsthand affected

We study the effects of contribution in relation of participants affected with MD. We can observe the average contribution and the evolution of contribution over round among participants affected and not affected. We observe significant differences between participants with MD and the rest of the collective. However, within the group of participants with MD we observe no significant differences in their contributions (see Fig.6.5). The most of participants with MD (43.6%) selected the maximum contribution (4), while the caregivers (46.5%) and non-caregivers (48.9%) mostly selected the fair contribution (2).

Effects of group composition

Each game was composed of participant with different roles in the ecosystems (affected with MD, caregivers and non-caregivers), we analyze if the distribution of participants with different roles has an effect on the collective contribution.

The evolution of the game (Fig.6.6), in which the most of participants are affected with MD and games with the most of participants are not affected, differs (t-test $p < 0.01$) in the first part of the game (rounds 1-5). Nonetheless average contribution in the last rounds has not significant differences (t-test $p > 0.05$).

The average contribution of participants with MD are 2.36 MU and 2.35 MU in the two contexts, with and without a majority of MD, is greater than the average contribution of the rest of collectives 2.09 MU and 2.10 MU (Fig.6.6). In average, the contribution based on role doesn't differ depending of the composition of the group.

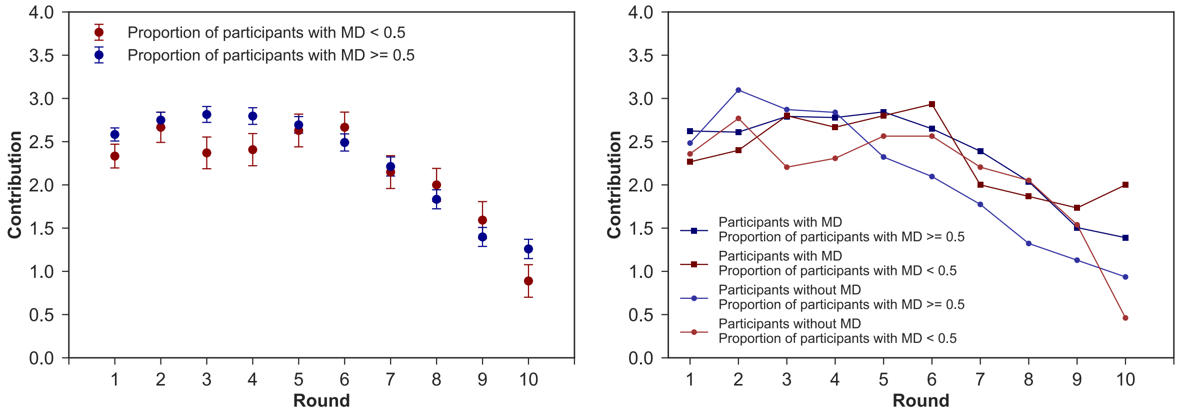

Figure 6.6: Contributions by group composition. Evolution of contributions (mean and standard error of the mean) over round according to the portion of participants with MD within groups. **Contributions by group composition and role.** Evolution of individual contributions over round of individuals with and without a mental condition by portion of individuals with MD within groups.

Inequalities

Participants affected with mental disorder tend to contribute more than the others in the collective-risk dilemma. This behaviour create inequalities, the most contribute the less earnings at the end of the game, we mesuare the inequalities created by this behaviour using the Gini index. We calculate the index in games with different proportions of participants with MD, and observe how in games with the most of participants affected with MD the inequalities increase, specially in the case with half of participants affected.

Earnings and Contributions

Final earnings include all participants (n=270). However, participants who did not contribute in two or more rounds, and had the computer contributing for them, did not get any profit. Their final earning is 0. The average (mean±sem, standard error of the mean) earnings of all subjects is 16.50 ± 0.53 MUs (see Fig. 6.7 for final earnings across groups).

Individual contributions per round (Fig.6.8) includes all participants (270). Yet, contributions in rounds where the computer selected for the participant were not included in the analysis. On the other hand, group contributions include both contributions made by the participants and computer selections.

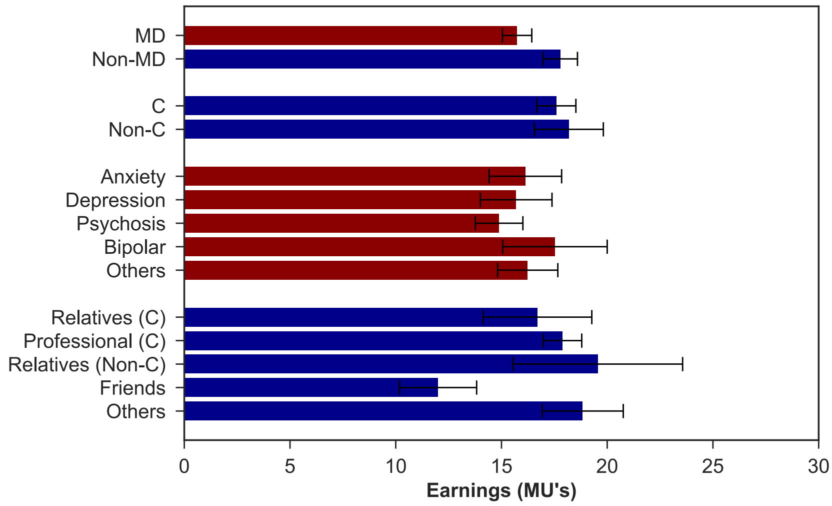

Figure 6.7: Earnings in Colletive-Risk Social Dilemma. Earnings (mean and standard error of the mean) by role in the ecosystem. We show results from participants with and without Mental Disorder condition (MD and Non-MD, respectively), caregivers and non caregivers (C and Non-C, respectively), MD individuals with different diagnosis, and finally other actors that may and may not be caregivers.

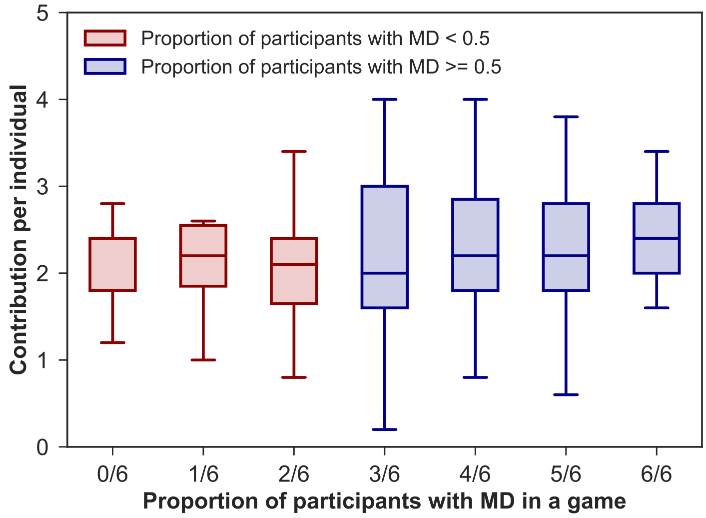

Figure 6.8: Individual contribution. There are not significant differences in individual contribution by groups composition (ANOVA, F: 0.371 p: 0.9).

Part IV

Concluding Remarks

Chapter 7

Conclusions and Perspectives

SUMMARY – In this last chapter we focus on the main conclusions extracted from the designed platforms and from the behavioral experiments explained and developed within this dissertation. We will also highlight the future lines of work and perspectives that derive from the research process, as well as the necessary steps that would assure an advancement of human behavior understanding and citizen engagement in science.

7.1 Conclusions

Nowadays there are new ways of doing science. Opening science to society, especially in disciplines that have a high social impact, is a must. We need a public debate about multiple scientific topics because science is dealing with issues that concern all of us. Citizen science (Bonney et al., 2014; Gura, 2013; Hand, 2010; Silvertown, 2009) can be a way to involve the general public in science. Introducing, to some extent, the public in the scientific process opens the door to questioning scientific work and to creating a society in which science is present in the public debate as a tool to address real social problems. A fortiori, the replicability crisis has meant that we have to rethink the way we do science, carrying out more studies of replication (Munafò et al., 2017; Nature, 2016a; McNutt, 2014) or even triangulation (Munafò and Smith, 2018), providing more statistical information in the experiments (Loken and Gelman, 2017) and offering the raw data in open repositories (Nature, 2014). In short, being more transparent doing science (Miguel et al., 2014) and enhancing the culture of open science (Nosek et al., 2015) within the scientific system itself.

In this dissertation we try to shed light on different topics. In the first place, we examine how to design participatory platforms that allow participation in science and that focus on the participants rather than on the scientific objective. From the need to move the experiments to the field, we work in the development of a flexible, modular and robust platform that encourages participation introducing human behaviour games. To throw light specifically on the platform, we present two experiments about decision-making with uncertainty and social dilemmas with different tensions and we focus on the robustness of the platform and the analysis of behavioural patterns. To study human behaviour in depth, we carried out two experiments with different objectives: the first analyzes how to solve a collective problem cooperatively, particularly if there are inequalities among the participants; and the second has members of the same community cooperate to solve a common problem and then evaluates if and which inequalities are created between them.

The platforms that have been created and presented in Chapter 2 and 3 allow, from two different perspectives, to bring science closer to the public based on citizen science practices, which are present in all the experiments that have been conducted. The platforms presented in here are fully designed to enhance participation in scientific activities. They are focused on the observation of natural surroundings, the understanding of human behaviour and the immersion in the scientific research process. In this dissertation we study the design process at different levels of participation and usage. It is essential to take into consideration and further improve the design procedures in different areas like the game mechanisms, the interaction between participants or the understanding of scientific concepts. In any case, the crucial idea is to put the experimenter at the inner center of the design. This applies to science in general, but especially to activities that are addressed at having an impact on people, as is the case of citizen science and particularly of crowdsourcing. In this vein, the need emerges to measure the impact beyond scientific contributions, which are expected in all scientific research, and look into the different contexts that flow from education to activism (Groulx et al., 2017; Jordan et al., 2011; Brossard et al., 2005). It is a requirement to ensure and stimulate all the potential outcomes (Bonney et al., 2015).

Natural Patterns, for instance, is a crowdsourcing platform whose storyline is the discovery and collection of natural patterns in the wild. We studied the design and participation in scientific activities mainly focusing on ideas such as science disposition, scientific thinking and user experience. The results presented are addressed in a study which was conducted to evaluate the performance of the design features of the application. In general, the participation in citizen science and crowdsourcing projects is concentrated in a few dedicated users (Sauermann and Franzoni, 2015), in the same way that in games there are a few who play the most (Pareto, 1896). In our study, most of the participants already had a high science disposition, so we focused on the analysis of the design features of the activity more than in the study of science disposition. The activities that are based on the observation of nature promote interaction with the participants' surroundings and, in general, the research process is well-represented in the storyline of Natural Patterns. The activities that require more involvement from the participants are more motivating than those that are simply proposed by the researchers, in the same way that the activities posed as a game. We really must adopt the gameplay mechanism to conduct science activities, and make them fun for the public (Ponti et al., 2018; Khelifa, 2016; Curtis, 2015; Bowser et al., 2013). From the study of Natural Patterns emerges the idea of designing scientific experiences by including the participant in most of the stages of the research process (Qaurooni et al., 2016; Bonney et al., 2015) rather than, as it has traditionally been done, focusing only on necessary scientific purposes.

Additionally, we present a platform for public participation in human behavior experiments which enables us to experiment in controlled non-laboratory environments using citizen science practices, such as those described in the work (Sagarra et al., 2016). The platform supports modular design of ad-hoc experiments, and the data collected are neither more nor less than that required by the design of the experiment and used exclusively for research purposes. Furthermore, the platform is designed as a game, in cases in which the experimental design permits it, and it collects metrics to study the decisions of participants from a qualitative point of view. The platform on human behavior experimentation and the way in which the experiments are performed encourage a part of society, traditionally not active in science, to engage in it. This can be observed in the demographic analysis of the experimental population far away from student samples with classic decision-making biases (Stoop, 2014; Exadaktylos et al., 2013; Falk et al., 2013; Carpenter et al., 2008). Experiments are carried out outside the laboratory, in the field, in real and controlled environments. Gamification techniques are used to enhance the engagement, the experience and the motivations of the participants, who are recruited randomly in-situ. Participation in this environments is much higher and diverse than in experiments which traditionally take place in laboratories, but less than the experiments conducted online with other platforms (Radford et al., 2016; Chen et al., 2016). However, in online experiments outside the lab the control of the subjects is not possible, especially in studies performed using systems such as mechanical turk, in which some aspects about the participants, reliability, data quality or labour conditions are unclear (Hara et al., 2017; Paolacci and Chandler, 2014; Goodman et al., 2013; Mason and Suri, 2012; Horton et al., 2011). In the Chapter 3 we delve into the details of the platform and how it overcomes these limitations.

The platform of human behavior experimentation has travelled a long journey. It has been used in more than ten experiments on decision making and social dilemmas in diverse environments and cities, always with high levels of participation and ensuring data quality and high impact publications. The data is collected adhoc for each study of human behavior, which makes it possible to reach a database with great granularity.

Proof of the performance of the human behaviour platform are the two experiments presented in Chapter 4. The first one, Mr.Banks, was carried out at the DAU (Barcelona), at Sonar+D (Barcelona) and at CAPS (Brussels). Diverse platform configurations made it possible to perform experiments robustly and securely. Simplifying, the experiment consists in making decisions based on different levels of information provided to participants to predict stock market movements. The results of the two replica experiments (Sonar+D and CAPS) were very similar to the original one in probability to choose up/down and in the strategies, even with samples which were one order of magnitude smaller.

The second experiment, Dr.Brain, was performed at the DAU (Barcelona) with 541 participants. The experiment provided that the participants played in groups of between 15-25 people in different combinations of the dyadic games represented in the TS plane, and did so for 13-18 rounds. This experiment helps to understand how we behave when confronted with different social dilemmas. The main results suggest that individual behaviour, when we face that set of dyadic games, can be described and reduced to a set of phenotypes: envious, pessimist, trustful and a group of undefined. It is important to emphasize that the phenotypes emerge in the data analysis using unsupervised learning algorithms, that is to say that no previous classifications had been made. This point is interesting because we can classify the participants from unstructured data. Clustering techniques are very useful when the data is complex and, as a result, it is difficult to find classification patterns manually.

Introducing machine learning algorithms to study social science is an interesting approach from different perspectives. Machine learning can identify patterns in unstructured data and classify them in clusters. When the patterns emerge, the results of the analysis can vary depending on the algorithm used, so their interpretation is in the hands of the data analyst. It is an important step toward drawing fair and unbiased conclusions, meaning we have to be aware of human factors that can influence both the process of drawing conclusions and the types of conclusions drawn. It is, for example, the case of the confirmation bias, typical of conventional analysis, in which the researcher directly answers a particular question, but it also occurs in unsupervised learning analysis by favouring information or reaching results that confirm one's preconceived hypotheses.

In the last two chapters we study human behavior focused on collective problems and how we face them collectively, in this case particularized on climate change. The collective-risk dilemma (Hagel et al., 2017; Jacquet et al., 2013; Milinski et al., 2011, 2008), a version of the public goods game with threshold, enables us to study this phenomenon.

Chapter 5 describes the experiment and the results in detail. In this experiment, besides straining individual and group interests, we focus on the inequalities that exist between the parties involved. Note that the most vulnerable, those who are most economically disadvantaged, relatively contribute more than those with a favorable economic position. Thus, it seems that the sense of justice is more developed in those who have a more precarious situation. Therefore, inequalities may lead to unexpected problems in climate change mitigation or analogous situations. Moreover, our participants held unambiguous responsibility over their actions, whereas climate change is a global problem with diffuse shared duties. In this sense, we have proven that a good general education is not the remedy to avoid inequality in contribution, but promoting collective attitudes, which has proven to be one factor underlying cooperation (Buchan et al., 2009), may be a better solution to make individuals assume their responsibilities.

Chapter 6 proposes the same experiment but without economic inequalities. All participants have the same economic conditions, however the experiments are composed of people with diverse roles in the mental health ecosystem, basically people affected by mental disorders, caregivers and relatives. Individuals with mental disorders are the ones who contribute the most to the public good: they make larger efforts towards reaching the collective goal, thus playing a leading role for the functioning of the ecosystem. This implies that other members of the group need a lower investments to reach the common target. It also unveils the vulnerability of individuals with a diagnosis of mental disorders. Given that our lab in the field experiment has been carried out in a fully socialized context, we point out that this approach can be applied to any similar initiatives focused on the analysis of a particular ecosystem interactions.

To conclude, some final ideas. We generate a lot of data online, which is useful to study plenty of patterns in a big scale. However, in order to study specific topics about human behaviour it is necessary to build a robust platform, such as the one presented here, with which to perform controlled human behaviour experiments (Vicens et al., 2018). Even more so if the idea is to move experimentation from the lab to the field, involving the public in scientific research (Sagarra et al., 2016). In that case, it is completely necessary to take special attention to the design of the experiment, in particular the experimental platform, in order to amplify the experience of the participants and to have an impact in their science disposition (Vicens et al., 2018). Furthermore, socio-technical complex systems, experimental game theory or computational social science are areas of study that contribute with evidence to topics that have a high impact in societal issues and that describe how we behave in situations of decision-making (Gutiérrez-Roig et al., 2016b), collective risk (Vicens et al., 2017) or personal-social conflicts (Poncela-Casasnovas et al., 2016). Also social dilemmas have an application in the study of the relationships between particular ecosystems, (Cigarini et al., 2018b) shedding light on the vulnerabilities (or strengths) of the ecosystem, but also on our society.

7.2 Perspectives

Science has always been on the periphery of society, while arts are at the center of culture. In the nineties, the concept of The Third Culture (Kelly, 1998; Brockman, 1996; Snow, 1959) and the cultural movements related with it opened a new way to communicate ideas to the public and introduced fresh air in the relationship between science and society. However, years afterwards, it would seem we have made little progress. Science is a part of a society's culture; however, for different reasons it has been relegated to a residual cultural position to which only those who are part of the scientific community, or even of the same discipline, have access. One way to aerate science is to actively involve society in it. This requires an effort in the development of scientific activities for non scientific communities, approaching and motivating participation to empower people in science. We need a scientifically empowered society to move forward, a society that embraces critical thinking, that questions everything and that seeks answers to the unknowns of the world.

In case that the significance of science in our society is still unclear, specifically in the topics that have been investigated in this dissertation, we only mention two examples to understand its impact in real-life situations relating to the study of human behavior and citizen science practices. The Obama Administration (2015b) issued an executive order called "*Executive Order - Using Behavioral Science Insights to Better Serve the American People*" in which it announces "*research findings from fields such as economics and behavioral psychology about how* *people make decisions and act on them -. can be used to design government policies to better serve the American people* ". In the same vein, empowering citizen science practices addressed to societal and scientific challenges (Obama Administration, 2015a) and promoting initiatives in the public sector (Jenn Gustetic et al., 2016).

In this dissertation we have studied experimentally how we face social dilemmas that put the individual and the collective interest in tension in full open environments, in the field, and also within a specific ecosystem, in order to observe the relationships and their equilibria. Although the initial premise is conducting experiments outside the laboratory in a controlled way, a clear step up is to create a completely stable and modular version for large-scale online experimentation. This opens up new possibilities to advance in a number of directions. We could, for instance, create an online platform point to study behavior in uncontrolled scenarios, to explore the influence of incentives, to prepare behavioural experiments with particular populations or cross-cultural experiments in real time. In general, we could reach diverse and large samples and conduct experiments with less infrastructure. There are studies using online platforms with relevant scientific contributions regarding the study of human behaviour (Stagnaro et al., 2017; Rand et al., 2014; Mason and Suri, 2012; Rand et al., 2012; Suri and Watts, 2011; Watts, 2007). However, online experimentation should be considered very rigorously because it may present some issues in verifying the identity of participants, the quality of data, the compensations, the privacy, and so on (Mason and Suri, 2012; Reips, 2002). But these issues are not only present at a scientific level; we need to work to avoid the inequalities and precarious work situations (Hara et al., 2017) that are generated in platforms already used to conduct research on human behavior or crowdsourcing activities. Therefore, it is completely necessary to address the performance of online experiences with rigorous protocols and to design them with the aim of engaging the participants in science beyond their participation as experimental subjects.

Social experimentation focusing on the study of social phenomena, the emergence of collective phenomena and behaviors in systems formed by individuals in interaction – cooperation, trust, reciprocity, social norms, etc. – is key to understand how human beings behave and interact with each other, especially when combined with simulation and theory. The study of behavioral traits has expanded from game theory to complex systems thanks to the computational study of these systems focused on behavioral models, introducing new approaches that expand the possibilities of knowledge in this field. This dissertation calls for experimentation by drawing general rules of human behavior, and pointing out the need to use experimental data to complement behavior models and simulations (Sánchez, 2018). However, experimentation can also aspire to have a direct impact in the decision-making process, for example in social policies. The primal insight of social experiments is the possibility to measure beyond opinions; to measure behaviors, inequalities and tensions. All this applied to social situations by means of carrying out framed experiments can work perfectly to measure the tensions of a society on specific topics.

The study of social phenomena, and specifically the study of human behavior, has people as its central object. Therefore, it is crucial to design experiments focusing on the participants, especially if the experiments are carried out in the wild, where the recruitment is totally random. To create an experimental environment that attracts people to participate in scientific research, it is essential that the participants obtain a return that goes beyond the economic incentive, some motivational factor like learning, recognition or, simply, enjoyment. The work to improve participation in experiments outside the laboratory is focused, above all, on the experimental platform, where it is necessary to introduce advances in HCI. With the objective of lightening the experiments and accelerating the learning curves, we should pay attention to the design of tutorials about the mechanism of the experiment and, mainly, make sure that the participants understand it perfectly before starting to make decisions autonomously. It is essential to focus on the design prior to conducting the experiment in order to assure the independence of the participants. In the case of experimentation in the wild, the experience expands beyond the platform to the physical environment. The "laboratory" in the field has to be a place where the participants feel a real interaction. In short, the design of the experiment in terms of environment and experimentation platform has to center on the participants and facilitate the protocols that ensure scientific rigor. In addition, besides the decision-making process, citizen science projects can maximize the participation of people before and after the experiment by promoting their involvement in co-designing and co-creating different phases of the research process. As a result, participation grows beyond the scientific community, and the concerns of participants are involved in the research process. Moreover, we can also create a community of participants who can follow all phases of the research process by means of the platform.

In the most recent experiments, that currently take place in the framework of other projects (e.g. STEM4Youth (Senabre et al., 2018)), new forms of participation in scientific experiences are applied. The participation as volunteer or experimental subject in the scientific experiment itself is still present and essential, but besides this there is citizen participation in the previous co-designing phases and subsequent stages of result analysis and interpretation (Senabre et al., 2018; Qaurooni et al., 2016; Tweddle et al., 2012; Shirk et al., 2012; Bonney et al., 2009). The co-design is especially crucial to guarantee an impact on the social needs of the community. The fundamental idea behind it is to try to respond to societal problems, so the participation of frontline affected subjects (citizens, associations, schools, municipalities) enriches the design process of the project and focuses the goals on real social problems. Co-designing/creating projects is a trend. In citizen science, it allows for the creation of activities around projects focused on outcomes such as learning, activism, enjoyment, and so on. Apparently, these practices need more resources than conventional initiatives because the creation of materials and the monitoring of the activities is essential. Thus, a great effort is devoted to the creation of a framework that makes the processes efficient. This line of work will allow to imbricate citizen science more profoundly in different layers of society.

Participatory science is an exercise in transparency that must be complemented with other open science actions. It is imperative that papers, especially those financed with public money, be available in open repositories, facilitating the access to scientific literature and helping in the divulgation of results. Also, the data of the experiments should be shared with the community, so that it does not only allow scientists to replicate studies, but also promotes that people with curiosity for data can play with them. The data should be treated with high security to preserve the privacy of the participants following the mechanisms of anonymization (Bowser et al., 2017). Therefore, it would be appropriate to share the code, the algorithms used to achieve results, in public repositories, so that anyone, even those with no knowledge of the subject but eager to learn, could execute the algorithms and perfectly reproduce the results of the studies. Undoubtedly, this would be a great advance for the scientific community in terms of accessibility and transparency, but also for society in general.

We have analyzed the experimental data using machine learning techniques to look for patterns in unstructured data by running unsupervised algorithms that emerge without a previous classification. The introduction of machine learning can also go beyond the analysis or the search for patterns. It would be the case of trying to predict behavior based on the data obtained. For example, in the phenotypes that have emerged from the study of the Chapter 4, where depending on which situations the participant cooperates or defects, so he/she belongs to one phenotype or another, we could look for a classifier or a model that allows, given a series of choices, the classification of the individual in a particular phenotype. Applying more robust and complex machine learning techniques to human behavior data can allow us to find patterns or predict behaviors as long as we have robust data. It is being applied in customer behavior, fight against terrorism or social policies, to name a few. The use of deep learning in the modeling and prediction of human behavior is a trend taking place especially in healthcare and social behavioral research (Gkotsis et al., 2017), but also in the study of models of cooperation and competition. If we enter in the field of artificial intelligence, natural languages and deep learning, we can use bots as a platform to conduct social experimentation and to study patterns that emerge not only from how we interact socially among us, but also from how we interact with machines that have artificial intelligence causing social phenomena (Stella et al., 2018; Ferrara, 2017) that would not come to light if the interaction were between individuals.

Today, many decisions are based on algorithms: news in social networks (Kramer et al., 2014), Internet searches (Google, 2018) or Netflix movie suggestions (Libby Plummer, 2017). But also whether we deserve a postdoctoral grant, and whether we are worth of being hired by a company. In short, algorithms make decisions that are transcendental in many cases, and we must analyze in depth whether they produce some kind of bias (Datta et al., 2015) or whether it is good to use them according to the circumstances (Matacic, 2018). Using machine learning algorithms to solve problems is more complicated than it seems, and they cannot be applied blindly to solve any kind of problem.

The data collected on proprietary platforms or devices has great potential for understanding us better, both individually and collectively. Therefore, it is important to note two things. First, that appropriately processed data provides us with valuable information about how we behave, allowing us to move forward along the path of knowledge, together with theory. Secondly, that the data generated for us in online platforms is held by diverse people – institutions, companies, researchers – with certain interests. Therefore we must be cautious with the data we generate, and demand transparency in the use of it. Giving up personal data could be seen as a donation, so we need more information about what data is being collected and the purposes behind the collection (Goodman and Flaxman, 2016).

We live deeply connected, receiving inputs from multiple sources. Scientific activity in this context cannot be carried out in isolation from the environment. More so when working with social data, which has the potential to represent the behavior of individuals and society. That's why a multidisciplinary approach is imperative , because the nature of the data forces different disciplines to cooperate in order to obtain the full potential of the knowledge encased in the data. Computational social science (Lazer et al., 2009; Conte et al., 2012; Mann, 2016) is an example of the need to work across disciplines in a natural way, and not do so in isolation. Technology allows us to generate, collect and process data that must be modeled, and then interpreted. In this apparently trivial process, many people could intercede: engineers, physicians or mathematics, data scientist, social scientist, and so on. A core of principles is needed to perform this research from multiples disciplines (Grimmer, 2015). Finally, the research should be more solution-oriented (Watts, 2017) and the advances generated should be addressed to society, to the people that directly or indirectly have facilitated the data (sometimes even the funds). Integrating the public in this process is interesting to boost the understanding of science and facilitate public scrutiny.

To synthesize, here are the main ideas and possibilities for the future based on the work done. To begin with, it would be interesting that the next steps in behavioral experimentation focused on specific social problems and on ecosystems with vulnerabilities, mainly, but not necessarily, human. Concretely for human behavior, a stimulating line of work would be the application of decision-making and social dilemmas experiments to study collective problems, especially in environments with inequalities. This would enable the application of measures that reduce tensions and therefore balance the ecosystem. Experimentation will allow better models and simulations to study how to apply these measures. Apart from the study of human behavior, the social phenomena that emerge from the interaction with machines (e.g. bots) and biological ecosystems could be examples of behaviors that affect us directly and which is necessary to understand.

In addition, the application of machine learning algorithms, and specifically deep learning, can help to find more sophisticated behavior patterns in the data collections that are extracted from the experiments and in data collected in other platforms. The experimental data collected in fully controlled situations and the data collected from platforms are intended to work together in order to improve the creation of realistic behavior models.

Open data policies can help in this sense, both public generic data and data devoted to scientific research. The positive trend to make science more open changes the way in which scientists, but also society, approach scientific knowledge. This movement works towards a situation in which open code, open data, open papers and open public participation (in science) are regular practices and essential for the future of science and society.

But to achieve this it is essential to promote citizen participation and connect people with science. This is a job that falls on scientists, even if it involves extra work or it may be uncomfortable to leave the community in which investigations normally circulate. All the advances in the aforementioned fields do not make any sense if they do not have a short/medium/long-term impact on society. Scientists can continue working in laboratories but the decisions that make changes happen take place outside them. It is necessary to position ourselves to face the challenges we have: what techniques do we use to reduce global warming, at the collective level and at the individual level; how we live in a safer world while maintaining the privacy of our data; how we manage virus epidemics or the migratory movements of refugees. From the perspective of science there are solutions to some of these problems, sometimes coming from very different approaches. To make this solutions a reality, however, in many cases there is a need for policies to be applied and for citizens to push for their application. Without informed citizens that are acquainted with the scientific proposals to solve a society's problems, it is not possible to advance. In the next few years I hope that science will open up and scientists will take their research closer to society, so we all can consider ourselves part of the scientific advances.

Bibliography

Abou Chakra, M. and A. Traulsen (2012). Evolutionary Dynamics of Strategic Behavior in a Collective-Risk Dilemma. *PLoS Computational Biology 8*(8).

Abou Chakra, M. and A. Traulsen (2014). Under high stakes and uncertainty the rich should lend the poor a helping hand. *Journal of Theoretical Biology 341*, 123–130.

Adam Pasick (2015). The magic that makes Spotify's Discover Weekly playlists so damn good.

Adams, J. (2012). Collaborations: The rise of research networks. *Nature 490*(7420), 335–336.

Adger, N. (2003). Social Capital, Collective Action, and Adaptation to Climate Change. *Economic Geography 79*(4), 387–404.

Agay, N., S. Kron, Z. Carmel, S. Mendlovic, and Y. Levkovitz (2008). Ultimatum bargaining behavior of people affected by schizophrenia. *Psychiatry Research 157*(1-3), 39–46.

Alanyali, M., H. S. Moat, and T. Preis (2013). Quantifying the relationship between financial news and the stock market. *Scientific Reports 3*.

Aledavood, T., A. M. Triana Hoyos, T. Alakörkkö, K. Kaski, J. Saramäki, E. Isometsä, and R. K. Darst (2017, jun). Data Collection for Mental Health Studies Through Digital Platforms: Requirements and Design of a Prototype. *JMIR Research Protocols 6*(6), e110.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders*.

Axelrod, R. and W. D. Hamilton (1981a). The evolution of cooperation. *Science (New York, N.Y.) 211*(4489), 1390–1396.

Axelrod, R. and W. D. Hamilton (1981b). The evolution of cooperation. *Science 211*(4489), 1390–1396.

Azar, C. and T. Sterner (1996). Discounting and distributional considerations in the context of globalwarming. *Ecological Economics 19*(2), 169–184.

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. *Processing 2*(3), 353.

Balliet, D., C. Parks, and J. Joireman (2009, jul). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes and Intergroup Relations 12*(4), 533–547.

Banerjee, A., M. S. Horn, and P. Davis (2016). Invasion of the Energy Monsters : A Family Board Game about Energy Consumption. *CHI Extended Abstracts on Human Factors in Computing Systems*, 1828–1834.

Barrett, S. and A. Dannenberg (2012). Climate negotiations under scientific uncertainty. *PNAS 109*(43), 17372–17376.

Bell, D. E., H. Raiffa, and A. Tversky (1988). *Decision making : descriptive, normative, and prescriptive interactions*. Cambridge University Press.

Bénabou, R. and J. Tirole (2003, jul). Intrinsic and extrinsic motivation. *Review of Economic Studies 70*(3), 489–520.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity and social history. *Games and Economic Behavior 10*, 122–142.

Beuscart, J.-S., K. Mellet, M. Trespeuch, T. Beauvisage, J.-S. Beuscart, M. Trespeuch, J. Borge-Holthoefer, Y. Moreno, and T. Yasseri (2016). *At the Crossroads: Lessons and Challenges in Computational Social Science*, Volume 4.

Blaga, O. M., L. Vasilescu, and R. M. Chereches (2017, jul). Use and effectiveness of behavioural economics in interventions for lifestyle risk factors of non-communicable diseases: a systematic review with policy implications. *Perspectives in Public Health*, 175791391772023.

Blut, C., A. Crespi, D. Mersch, L. Keller, L. Zhao, M. Kollmann, B. Schellscheidt, C. Fülber, and M. Beye (2017, dec). Automated computer-based detection of encounter behaviours in groups of honeybees. *Scientific Reports 7*(1), 17663.

Bohnet, I. and B. S. Frey (1999). Social Distance and other-regarding Behavior in Dictator Games: Comment. *Source: The American Economic Review 89*(1), 335–339.

Bonney, R., H. Ballard, R. Jordan, E. McCallie, T. Phillips, J. Shirk, and C. C. Wilderman (2009). Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education Research : Defining the Field and Science Education. *Science Education*.

Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience 59*(11), 977–984.

Bonney, R., T. B. Phillips, H. L. Ballard, and J. W. Enck (2015). Can citizen science enhance public understanding of science? *Public Understanding of Science 25*(1), 2–16.

Bonney, R., J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish (2014). Citizen science: Next steps for citizen science. *Science 343*(6178), 1436–1437.

Bornmann, L. and R. Mutz (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology 66*(11), 2215–2222.

Bowser, A., D. Hansen, J. Preece, Y. He, C. Boston, L. Gunnell, and M. Reid (2013). Using gamification to inspire new citizen science volunteers. *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, 18–25.

Bowser, A. and L. Shanley (2013). New visions in citizen science. *Case Study Series 3*, 1–53.

Bowser, A., K. Shilton, E. Warrick, and J. Preece (2017). Accounting for privacy in citizen science: Ethical research in a context of openness Volunteer Motivation. *Proceedings of CSCW 2017*, 2124–2136.

Brabham, D. C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies 14*(1), 75–90.

Brandon Keim (2017). Alan Turing's Patterns in Nature, and Beyond.

Brockman, J. (1996). *The third culture.* Simon & Schuster.

Brossard, D., B. Lewenstein, and R. E. Bonney (2005). Scientific knowledge and attitude change : The impact of a citizen science project. *International Journal of Science Education 27*(July), 1099–1121.

Brown, A., P. Franken, S. Bonner, N. Dolezal, and J. Moross (2016). Safecast: successful citizen-science for radiation measurement and communication after Fukushima. *Journal of Radiological Protection 36*(2), S82–S101.

Brülhart, M. and J. C. Usunier (2012). Does the trust game measure trust? *Economics Letters 115*(1), 20–23.

Buchan, N. R., G. Grimalda, R. Wilson, M. Brewer, E. Fatas, and M. Foddy (2009, mar). Globalization and human cooperation. *Proceedings of the National Academy of Sciences of the United States of America 106*(11), 4138–42.

Bulkeley, H., G. A. Edwards, and S. Fuller (2014, mar). Contesting climate justice in the city: Examining politics and practice in urban climate change experiments. *Global Environmental Change 25*, 31–40.

Burgess, M., E. Adar, and M. Cafarella (2016, may). Link-Prediction Enhanced Consensus Clustering for Complex Networks. *PLOS ONE 11*(5), e0153384.

Burton-Chellew, M. N., R. M. May, and S. A. West (2013). Combined inequality in wealth and risk leads to disaster in the climate change game. *Climatic Change 120*(4), 815–830.

Bynum, A., R. Kline, and O. Smirnov (2016). Passive non-participation versus strategic defection in a collective risk social dilemma. *Journal of Theoretical Politics 28*(1), 138–158.

CAISE Inquiry Group (2009). Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education Research : Defining the Field and Science Education. *Science Education* (July).

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction.*

Carpenter, J., C. Connolly, and C. K. Myers (2008). Altruistic behavior in a representative dictator experiment. *Experimental Economics 11*(3), 282–298.

Ceccaroni, L., A. Bowser, and P. Brenton (2017). Civic Education and Citizen Science:. *Analyzing the Role of Citizen Science in Modern Research*, 1–23.

Cechanowicz, J., C. Gutwin, B. Brownell, and L. Goodfellow (2013). Effects of gamification on participation and data quality in a real-world market research domain. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications - Gamification '13*, New York, New York, USA, pp. 58–65. ACM Press.

Cesarini, D., C. T. Dawes, M. Johannesson, P. Lichtenstein, and B. Wallace (2009). Experimental game theory and behavior genetics. *Annals of the New York Academy of Sciences 1167*, 66–75.

Chan, K. S., S. Mestelman, R. Moir, and R. A. Muller (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics 2*(1), 5–30.

Chang, A. C. and P. Li (2015). Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". *Finance and Economics Discussion Series 2015*(83), 1–26.

Charrad, M., N. Ghazzali, V. Boiteau, and A. Niknafs (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software 61*(6).

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Chen, H., W. Chung, J. Xu, G. Wang, Y. Qin, and M. Chau (2004, apr). Crime data mining: a general framework and some examples. *Computer 37*(4), 50–56.

Chen, J. (2007, apr). Flow in games (and everything else). *Communications of the ACM 50*(4), 31.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002). Hardnose the dictator. *American Economic Review 92*(4), 1218–1221.

Cherry, T. L., S. Kroll, and J. F. Shogren (2005). The impact of endowment heterogeneity and origin on public good contributions: Evidence from the lab. *Journal of Economic Behavior and Organization 57*(3), 357–365.

Cigarini, A., J. Vicens, J. Duch, A. Sánchez, and J. Perelló (2018a, feb). Dataset – Quantitative account of social interactions in a mental health care ecosystem: cooperation, trust and collective action.

Cigarini, A., J. Vicens, J. Duch, A. Sánchez, and J. Perelló (2018b, dec). Quantitative account of social interactions in a mental health care ecosystem: cooperation, trust and collective action. *Scientific Reports 8*(1), 3794.

Cioffi-Revilla, C. (2010). Computational social science.

Cohn, J. P. (2008). Citizen Science: Can Volunteers Do Real Research? *BioScience 58*(3), 192.

Consortium, S. (2013). *Green paper on citizen science. Citizen Science for Europe. Towards a better society of empowered citizens and enhanced research.* Brussels.

Conte, R., N. Gilbert, G. Bonelli, C. Cioffi-Revilla, G. Deffuant, J. Kertesz, V. Loreto, S. Moat, J. P. Nadal, A. Sanchez, A. Nowak, A. Flache, M. San Miguel, and D. Helbing (2012). Manifesto of computational social science. *European Physical Journal: Special Topics 214*(1), 325–346.

Cooper, C., J. Dickinson, T. Phillips, and R. Bonney (2007). Citizen Science as a Tool for Conservation in Residential Ecosystems - Citizen science Cooper et al 2007.pdf. *Ecology and Society 12*(2), Thursday, December 13, 2007.

Cooper, C. B., J. Shirk, and B. Zuckerberg (2014). The invisible prevalence of citizen science in global research: Migratory birds and climate change. *PLoS ONE 9*(9).

Cooper, S., F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, and Z. Popović (2010). Predicting Protein Structures with a Multiplayer Online Game. *Nature 466*(August), 756–760.

Cornell Lab. Bioacoustics Research Program. (2018). Right Whale Listening Network.

Crall, A. W., R. Jordan, K. Holfelder, G. J. Newman, J. Graham, and D. M. Waller (2013). The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public understanding of science (Bristol, England) 22*(6), 745–64.

Crawford, V. P. (2002). Introduction to experimental game theory. *Journal of Economic Theory 104*(1), 1–15.

Crowston, K. and N. R. Prestopnik (2013). Motivation and data quality in a citizen science game: A design science evaluation. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 450–459.

Csukly, G., P. Polgár, L. Tombor, J. Réthelyi, and S. Kéri (2011). Are patients with schizophrenia rational maximizers? Evidence from an ultimatum game study. *Psychiatry Research 187*(1-2), 11–17.

Curtis, V. (2015). Motivation to Participate in an Online Citizen Science Game: A Study of Foldit. *Science Communication 37*(6), 723–746.

Danielsen, F., N. D. Burgess, A. Balmford, P. F. Donald, M. Funder, J. P. G. Jones, P. Alviola, D. S. Balete, T. Blomley, J. Brashares, B. Child, M. Enghoff, J. Fjelds, S. Holt, H. Hebertz, A. E. Jensen, P. M. Jensen, J. Massao, M. M. Mendoza, Y. Ngaga, M. K. Poulsen, R. Rueda, M. Sam, T. Skielboe, G. Stuart-Hill, E. Topp-Jorgensen, and D. Yonten (2009). Local participation in natural resource monitoring: A characterization of approaches. *Conservation Biology 23*(1), 31–42.

Dannenberg, A., A. Löschel, G. Paolacci, C. Reif, and A. Tavoni (2015). On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds. *Environmental and Resource Economics 61*(3), 365–383.

Datta, A., M. C. Tschantz, and A. Datta (2015). Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination. *Proceedings on Privacy Enhancing Technologies 2015*(1), 92–112.

Davies, D. L. and D. W. Bouldin (1979, apr). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-1*(2), 224–227.

Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology 31*(2), 169–193.

de Oliveira Roque, F. (2016, sep). Field studies: Could Pokémon Go boost birding? *Nature 537*(7618), 34–34.

Deci, E. L. and R. M. Ryan (1985). Intrinsic motivation and self-determination in human behavior: Springer Science & Business Media.

Delfanti, A. (2010). Users and peers. from citizen science to P2P science. *Journal of Science Communication 9*(1), 1–5.

Denis, G. and P. Jouvelot (2005). Motivation-driven educational game design. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '05*, New York, New York, USA, pp. 462–465. ACM Press.

Dennig, F., M. B. Budolfson, M. Fleurbaey, A. Siebert, and R. H. Socolow (2015). Inequality, climate impacts on the future poor, and carbon prices. *Proceedings of the National Academy of Sciences 112*(52), 15827–15832.

Derrick Harris (2012). The one-night stand, quantified and visualized by Uber.

Deterding, S. (2012, jul). Gamification: designing for motivation. *interactions 19*(4), 14.

Deterding, S., D. Dixon, R. Khaled, and L. Nacke (2011). From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, 9–11.

D'Hondt, E., M. Stevens, and A. Jacobs (2013). Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. *Pervasive and Mobile Computing 9*(5), 681–694.

Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K. Purcell (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment 10*(6), 291–297.

Dieleman, S., K. W. Willett, and J. Dambre (2015, mar). Rotation-invariant convolutional neural networks for galaxy morphology prediction. *arXiv*.

Dorward, L. J., J. C. Mittermeier, C. Sandbrook, and F. Spooner (2017, jan). Pokémon Go: Benefits, Costs, and Lessons for the Conservation Movement. *Conservation Letters 10*(1), 160–165.

Dutta, P. K. and R. Radner (2009). A strategic analysis of global warming: Theory and some numbers. *Journal of Economic Behavior and Organization 71*(2), 187–209.

Editorial (2015). Rise of the citizen scientist. *Nature 524*(7565), 265–265.

Egnor, S. R. and K. Branson (2016, jul). Computational Analysis of Behavior. *Annual Review of Neuroscience 39*(1), 217–236.

Einav, L. and J. Levin (2014). Economics in the age of big data. *Science 346*(6210).

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik 322*(10), 891–921.

Emily Matchar (2017). AI Plant and Animal Identification Helps Us All Be Citizen Scientists. *Smithsonian Magazine*.

Eriksson, M. (2011). Social capital and health–implications for health promotion. *Global health action 4*, 5611.

Exadaktylos, F., A. M. Espín, and P. Brañas-Garza (2013). Experimental subjects are not different. *Scientific Reports 3*(1213), 1–6.

Falk, A., E. Fehr, and U. Fischbacher (2003). On the nature of fair behavior. *Economic Inquiry 41*(1), 20–26.

Falk, A. and J. J. Heckman (2009). Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science 326*(5952), 535–539.

Falk, A., S. Meier, and C. Zehnder (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association 11*(4), 839–852.

Farmer, J. D. and A. W. Lo (1999, aug). Frontiers of finance: Evolution and efficient markets. *Proceedings of the National Academy of Sciences 96*(18), 9991–9992.

Fehr, E. and U. Fischbacher (2004). Third-party punishment and social norms. *Evolution and Human Behavior 25*(2), 63–87.

Fehr, E. and S. Gächter (2002). Altruistic punishment in humans. *Nature 415*(6868), 137–140.

Fehr, E. and J. A. List (2004). The Hidden Costs and Returns of IncentivesTrust and Trustworthiness among Ceos. *Journal of the European Economic Association 2*(5), 743–771.

Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics 114*(3), 817–868.

Ferrara, E. (2017, aug). Measuring social spam and the effect of bots on information diffusion in social media. *arXiv*.

Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are People Conditionally Cooperative ? Evidence from a Public Goods Experiment Duh Shrsoh Frqglwlrqdoo | FrrshudwlyhB Hylghqfh iurp d Sxeolf Jrrgv H { shulphqw. *Economics Letters 71*(3), 397–404.

Franzoni, C. and H. Sauermann (2014). Crowd science: The organization of scientific research in open collaborative projects. *Research Policy 43*(1), 1–20.

Freytag, A., W. Güth, H. Koppel, and L. Wangler (2014). Is regulation by milestones efficiency enhancing? An experimental study of environmental protection. *European Journal of Political Economy 33*, 71–84.

Gao, H., G. Barbier, and R. Goolsby (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems 26*(3), 10–14.

Gentile, D. A., C. A. Anderson, S. Yukawa, N. Ihori, M. Saleem, L. K. Ming, A. Shibuya, A. K. Liau, A. Khoo, B. J. Bushman, L. Rowell Huesmann, and A. Sakamoto (2009, jun). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. *Personality and Social Psychology Bulletin 35*(6), 752–763.

Giles, J. (2012, aug). Computational social science: Making the links. *Nature 488*(7412), 448–450.

Gintis, H. (2000). Beyond Homo economicus: Evidence from experimental economics. *Ecological Economics 35*(3), 311–322.

Gkotsis, G., A. Oellrich, S. Velupillai, M. Liakata, T. J. P. Hubbard, R. J. B. Dobson, and R. Dutta (2017, mar). Characterisation of mental health conditions in social media using Informed Deep Learning. *Scientific Reports 7*, 45141.

González-Bailón, S., J. Borge-Holthoefer, A. Rivero, and Y. Moreno (2011). The Dynamics of Protest Recruitment through an Online Network.

Goodman, B. and S. Flaxman (2016). European Union regulations on algorithmic decision-making and a "right to explanation". pp. 1–9.

Goodman, J. K., C. E. Cryder, and A. Cheema (2013, jul). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making 26*(3), 213–224.

Google (2018). How Google Search Works.

Gradin, V. B., A. Pérez, J. A. MacFarlane, I. Cavin, G. Waiter, J. Engelmann, B. Dritschel, A. Pomi, K. Matthews, and J. D. Steele (2015). Abnormal brain responses to social fairness in depression: an fMRI study using the Ultimatum Game. *Psychological medicine 45*(6), 1241–51.

Greene, M. (2007). The demise of the lone author. *Nature 450*(7173), 1165.

Greitemeyer, T. and S. Osswald (2010). Effects of Prosocial Video Games on Prosocial Behavior. *Journal of Personality and Social Psychology 98*(2), 211–221.

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. In *PS - Political Science and Politics*, Volume 48, pp. 80–83.

Groulx, M., M. C. Brisbois, C. J. Lemieux, A. Winegardner, and L. A. Fishback (2017, feb). A Role for Nature-Based Citizen Science in Promoting Individual and Collective Climate Change Action? A Systematic Review of Learning Outcomes. *Science Communication 39*(1), 45–76.

Guala, F. and L. Mittone (2010). Paradigmatic experiments: The Dictator Game. *Journal of Socio-Economics 39*(5), 578–584.

Gura, T. (2013). Citizen science: amateur experts. *Nature 496*(7444), 259–261.

Guroglu, B., W. van den Bos, S. A. R. B. Rombouts, and E. A. Crone (2010). Unfair? It depends: Neural correlates of fairness in social context. *Social Cognitive and Affective Neuroscience 5*(4), 414–423.

Gutierrez-Roig, M. (2015). *Experiments and Models for Human Decision-Making: Social Dilemmas, Pedestrian Movements and Financial Markets.* Ph. D. thesis, Universitat de Barcelona.

Gutiérrez-Roig, M., C. Segura, J. Duch, and J. Perelló (2016a). Dataset – Mr.Banks Experiment 2013.

Gutiérrez-Roig, M., C. Segura, J. Duch, and J. Perelló (2016b). Market imitation and win-stay lose-shift strategies emerge as unintended patterns in market direction guesses. *PLoS ONE 11*(8).

Hagel, K., M. A. Chakra, B. Bauer, and A. Traulsen (2016). Which risk scenarios can drive the emergence of costly cooperation? *Scientific Reports 6*.

Hagel, K., M. Milinski, and J. Marotzke (2017). The level of climate-change mitigation depends on how humans assess the risk arising from missing the 2 degrees target. *Palgrave Communications 3*, 17027.

Haklay, M. (2013). Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, Volume 9789400745, pp. 105–122.

Haklay, M. (2015). Citizen Science and Policy: A European Perspective. *Common Labs. Case Study Series. 4*, 76.

Hand, E. (2010). People Power. *Nature 466*(August), 685–687.

Hara, K., A. Adams, K. Milland, S. Savage, C. Callison-Burch, and J. Bigham (2017, dec). A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. *arXiv*.

Harburg, E., Y. Kim, E. Gerber, and H. Zhang (2015). CrowdFound: A Mobile Crowdsourcing System to Find Lost Items On-the-Go. *Extended Abstracts of the ACM CHI'15 Conference on Human Factors in Computing Systems 2*, 1537–1542.

Hauert, C., S. De Monte, J. Hofbauer, and K. Sigmund (2002). Replicator dynamics for optional public good games. *Journal of Theoretical Biology 218*(2), 187–194.

Heyes, C. (2011). Automatic Imitation. *Psychological Bulletin 137*(3), 463–483.

Hilbe, C., M. Abou Chakra, P. M. Altrock, and A. Traulsen (2013). The Evolution of Strategic Timing in Collective-Risk Dilemmas. *PLoS ONE 8*(6).

Holt, C. (2018). Vecon Lab: http://veconlab.econ.virginia.edu/.

Horn, M. S. (2014). Beyond Video Games for Social Change. *Interactions*, 66–68.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics 14*(3), 399–425.

Howe, J. (2006). The Rise of Crowdsourcing. *Wired Magazine 14*(06), 1–5.

Huang, Z. (1997). Clustering large data sets with mixed numeric and categorical values. *Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining,(PAKDD)*, 21–34.

Hunicke, R., M. LeBlanc, and R. Zubek (2004). MDA: A Formal Approach to Game Design and Game Research. *Workshop on Challenges in Game AI*, 1–4.

Imran, M., P. Meier, C. Castillo, A. Lesa, and M. Garcia Herranz (2016). Enabling Digital Health by Automatic Classification of Short Messages. *Proceedings of the 6th International Conference on Digital Health Conference - DH '16*, 61–65.

INaturalist (2017). iNaturalist Computer Vision Explorations.

Insel, T. R., P. Y. Collins, and S. E. Hyman (2015). Darkness invisible: The hidden global costs of mental illness.

JABONERO, V. (1953, nov). [On the synapses in the sympathetic ganglia]. *Archivo espanol de morfologia 10*(35-38), 109–27.

Jacquet, J., K. Hagel, C. Hauert, J. Marotzke, T. Röhl, and M. Milinski (2013). Intra- and intergenerational discounting in the climate game. *Nature Climate Change 3*(12), 1025–1028.

Jenn Gustetic, Kelly Olson, and Anne Bowser (2016). Collaboration Gives Federal Government Citizen Science and Crowdsourcing a New Home on the Web | whitehouse.gov.

Jennett, C., L. Kloetzer, D. Schneider, I. Iacovides, A. L. Cox, M. Gold, B. Fuchs, A. Eveleigh, K. Mathieu, Z. Ajani, and Y. Talsi (2016). Motivations, learning and creativity in online citizen science. *Journal of Science Communication 15*(3).

Jess McNally (2017). Earth's Most Stunning Natural Fractal Patterns.

Jordan, R. C., S. A. Gray, D. V. Howe, W. R. Brooks, and J. G. Ehrenfeld (2011). Knowledge Gain and Behavioral Change in Citizen-Science Programs. *Conservation Biology 25*(6), 1148–1154.

Julie Muncy (2016). Making AI Play Lots of Videogames Could Be Huge (No, Seriously).

Kang, P. and S. Cho (2009). K-means clustering seeds initialization based on centrality, sparsity, and isotropy. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 5788 LNCS, pp. 109–117. Springer, Berlin, Heidelberg.

Kelly, K. (1998, feb). The Third Culture. *Science 279*(5353), 992–993.

Kennard, F. (2015). *Thought Experiments: Popular Thought Experiments in Philosophy, Physics, Ethics, Computer Science & Mathematics.*

Khelifa, R. (2016). Lure gamers into citizen science.

King-Casas, B. and P. H. Chiu (2012). Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biological Psychiatry 72*(2), 119–125.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *Appears in the International Joint Conference on Articial Intelligence (IJCAI)*, Volume 118, pp. 1–7.

Kollock, P. (1998). Social dilemmas: The Anatomy of Cooperation. *Annu. Rev. Sociol 24*(1998), 183–214.

Kramer, A. D. I., J. E. Guillory, and J. T. Hancock (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences 111*(29), 10779–10779.

Krieger, A. M. and P. E. Green (1999, jan). A Generalized Rand-Index Method for Consensus Clustering of Separate Partitions of the Same Data Base. *Journal of Classification 16*(1), 63–89.

Kullenberg, C. and D. Kasperowski (2016). What is citizen science? - A scientometric meta-analysis. *PLoS ONE 11*(1), 1–16.

Lancichinetti, A. and S. Fortunato (2012, dec). Consensus clustering in complex networks. *Scientific Reports 2*(1), 336.

Lazer, D., A. Pentland, L. Adamic, S. Aral, A.-L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne (2009). Computational social science. *Science 323*(5915), 721–3.

Lee, T. Y., C. Dugan, W. Geyer, T. Ratchford, J. Rasmussen, N. S. Shami, and S. Lupushor (2013). Experiments on motivational feedback for crowdsourced workers. *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, 341–350.

Leshner, A. I. (2015). Bridging the opinion gap. *Science 347*(6221), 459.

Levitt, S. D. and J. A. List (2007a). Viewpoint: On the generalizability of lab behaviour to the field.

Levitt, S. D. and J. A. List (2007b). What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real? *Journal of Economic Perspectives 21*(2), 153–174.

Levitt, S. D. and J. A. List (2008). Homo economicus Evolves. *Science 319*(5865), 909–910.

Li, A., J. Walling, S. Ahn, Y. Kotliarov, Q. Su, M. Quezado, J. C. Oberholtzer, J. Park, J. C. Zenklusen, and H. A. Fine (2009). Unsupervised Analysis of Transcriptomic Profiles Reveals Six Glioma Subtypes. *Cancer Res 69*(5), 2091–9.

Libby Plummer (2017). This is how Netflix's top-secret recommendation system works.

Licht, A. N. (1999). Games Commissions Play : 2x2 Games of International Securities Regulation. *The Yale Journal of International Law 24*(61), 61–125.

Lintott, C. J., K. Schawinski, A. Slosar, K. Land, S. Bamford, D. Thomas, M. J. Raddick, R. C. Nichol, A. Szalay, D. Andreescu, P. Murray, and J. Vandenberg (2008). Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society 389*(3), 1179–1189.

List, J. A. (2009). An introduction to field experiments in economics. *Journal of Economic Behavior and Organization 70*(3), 439–442.

Lock, E. F. and D. B. Dunson (2013, oct). Bayesian consensus clustering. *Bioinformatics 29*(20), 2610–2616.

Loken, E. and A. Gelman (2017, feb). Measurement error and the replication crisis. *Science (New York, N.Y.) 355*(6325), 584–585.

MacKay, D. J. C. (2005). *Information Theory, Inference, and Learning Algorithms*, Volume 100. Cambridge University Press.

MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. *5th Berkeley Symposium on Mathematical Statistics and Probability 1967 1*(233), 281–297.

Malone, T. W., R. Laubacher, and C. Dellarocas (2010). The collective intelligence genome. *IEEE Engineering Management Review 38*(3), 38.

Mann, A. (2016). Core Concept: Computational social science. *Proceedings of the National Academy of Sciences 113*(3), 468–470.

Mason, W. and S. Suri (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods 44*(1), 1–23.

Massung, E., D. Coyle, K. F. Cater, M. Jay, and C. Preist (2013). Using crowdsourcing to support pro-environmental community activism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, pp. 371–380.

Matacic, C. (2018). Are algorithms good judges? *Science 359*(6373), 263–263.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*, Volume 13. Cambridge University Press.

McNutt, M. (2014). Reproducibility. *Science 343*, 229.

Mendelsohn, R., A. Dinar, and L. Williams (2006). The distributional impact of climate change on rich and poor countries. *Environment and Development Economics 11*(02), 159.

Miguel, E., C. Camerer, K. Casey, J. Cohen, K. M. Esterling, A. Gerber, R. Glennerster, D. P. Green, M. Humphreys, G. Imbens, D. Laitin, T. Madon, L. Nelson, B. A. Nosek, M. Petersen, R. Sedlmayr, J. P. Simmons, U. Simonsohn, and M. Van Der Laan (2014). Promoting transparency in social science research.

Milinski, M., C. Hilbe, D. Semmann, R. Sommerfeld, and J. Marotzke (2016). Humans choose representatives who enforce cooperation in social dilemmas through extortion. *Nature Communications 7*, 10915.

Milinski, M., T. Röhl, and J. Marotzke (2011). Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change 109*(3-4), 807–814.

Milinski, M., R. D. Sommerfeld, H. J. Krambeck, F. A. Reed, and J. Marotzke (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *PNAS 105*(7), 2291–2294.

Miller-Rushing, A., R. Primack, and R. Bonney (2012). The history of public participation in ecological research. *Frontiers in Ecology and the Environment 10*(6), 285–290.

Mokros, A., B. Menner, H. Eisenbarth, G. W. Alpers, K. W. Lange, and M. Osterheider (2008). Diminished Cooperativeness of Psychopaths in a Prisoner's Dilemma Game Yields Higher Rewards. *Journal of Abnormal Psychology 117*(2), 406–413.

Monti, S., P. Tamayo, J. Mesirov, and T. Golub (2003). Consensus Clustering. *Machine Learning 52*(1), 1–34.

Morris, B. J., S. Croker, C. Zimmerman, D. Gill, and C. Romig (2013, sep). Gaming science: The "Gamification" of scientific thinking.

Morschheuser, B., J. Hamari, and J. Koivisto (2016a). Gamification in crowdsourcing: A review. *Proceedings of the Annual Hawaii International Conference on System Sciences 2016-March*, 4375–4384.

Morschheuser, B., J. Hamari, and J. Koivisto (2016b). Gamification in crowdsourcing: A review. *Proceedings of the Annual Hawaii International Conference on System Sciences 2016-March*, 4375–4384.

Munafò, M. R., B. A. Nosek, D. V. M. Bishop, K. S. Button, C. D. Chambers, N. Percie du Sert, U. Simonsohn, E.-J. Wagenmakers, J. J. Ware, and J. P. A. Ioannidis (2017). A manifesto for reproducible science. *Nature Human Behaviour 1*(1), 0021.

Munafò, M. R. and G. D. Smith (2018). Repeating experiments is not enough. *Nature 553*, 399–401.

Murtagh, F. and P. Legendre (2011, nov). Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. *Journal of Classification 31*(3), 274–295.

Museum of Science and Industry (2017). Explore the Patterns.

Myerson, R. B. (1997). *Game Theory: Analysis of Conflict.*

NASA (2017a). GLOBE Observer.

NASA (2017b). GLOBE Program.

NASA (2017c). iSeeChange.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences 36*(1), 48–49.

National Audubon Society (2017). Christmas Bird Count.

National Geographic (2017). Patterns in Nature.

Nature (2014, nov). Data-access practices strengthened. *Nature 515*(7527), 312.

Nature (2016a, aug). Go forth and replicate! *Nature 536*(7617), 373–373.

Nature (2016b, jul). Gotta name them all: how Pokémon can transform taxonomy. *Nature 535*(7612), 323–324.

Nature Editorial (2017a). Announcement: Transparency upgrade for Nature journals. *Nature 543*(7645), 288.

Nature Editorial (2017b, jan). Why researchers should resolve to engage in 2017. *Nature 541*(7635), 5–5.

NGO (2016). El Mussol.

Norheim-Hagtun, I. and P. Meier (2010). Crowdsourcing for Crisis Mapping in Haiti. *Innovations: Technology, Governance, Globalization 5*, 81–89.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni (2015). Promoting an open research culture. *Science 348*(6242).

Nov, O., O. Arazy, and D. Anderson (2011). Dusting for science: Motivation and participation of digital citizen science volunteers. *iConference*, 68–74.

Nowak, M. and K. Sigmund (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature 364*(6432), 56–58.

Nowak, M. A. (2006a). *Evolutionary Dynamics.* Harvard University Press.

Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science 314*(5805), 1560–1563.

Obama Administration (2015a). Accelerating Citizen Science and Crowdsourcing to Address Societal and Scientific Challenges.

Obama Administration (2015b). Executive Order No. 13707 – Using Behavioral Science Insights to Better Serve the American People.

Olson, M. (1971). *The logic of collective action : public goods and the theory of groups.* Harvard University Press.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science (New York, N.Y.) 349*(6251), aac4716.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist 17*(11), 776–783.

Oxford University Press (2014). *Oxford English Dictionary.* Oxford University Press.

Pachauri, R. and L. Meyer (2014). *Climate Change 2014: Synthesis Report. Contribution ofWorking Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC, Geneva, Switzerland.

Paolacci, G. and J. Chandler (2014, jun). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science 23*(3), 184–188.

Pareto, V. (1896). *Cours D'économie politique.*

Patel, J., S. Shah, P. Thakkar, and K. Kotecha (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications 42*(4), 2162–2172.

PBL Netherlands Environmental Assessment Agency (2016). Trends in Global CO_2 Emissions: 2016 Report. Technical report.

Perelló, J., N. Ferran-Ferrer, S. Ferré, T. Pou, and I. Bonhoure (2018). High motivation and relevant scientific competencies through the introduction of citizen science at Secondary schools:: An assessment using a rubric model.

Peysakhovich, A., M. a. Nowak, and D. G. Rand (2014). Humans Display a 'Cooperative Phenotype' that is Domain General and Temporally Stable. *Nature communications 5*(September), 1–8.

Pham, D. T., S. S. Dimov, and C. D. Nguyen (2005). Selection of K in K-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 219*(1), 103–119.

Poncela-Casasnovas, J., M. Gutiérrez-Roig, C. Gracia-Lázaro, J. Vicens, J. Gómez-Gardeñes, J. Perelló, Y. Moreno, J. Duch, and A. Sánchez (2016). Humans display a reduced set of consistent behavioral phenotypes in dyadic games. *Science advances 2*(8), e1600451.

Poncela-Casasnovas, J., M. Gutiérrez-Roig, C. Gracia-Lázaro, J. Vicens, J. Gómez-Gardeñes, J. Perelló, Y. Moreno, J. Duch, and A. Sánchez (2017, dec). Dataset – Humans display a reduced set of consistent behavioral phenotypes in dyadic games.

Ponti, M., T. Hillman, C. Kullenberg, and D. Kasperowski (2018). Getting it Right or Being Top Rank: Games in Citizen Science. *Citizen Science: Theory and Practice 3*(1), 1–12.

Pratt, J. W., H. Raiffa, and R. Schlaifer (1964, jun). The Foundations of Decision under Uncertainty: An Elementary Exposition. *Journal of the American Statistical Association 59*(306), 353–375.

Preece, J. (2016). Citizen Science: New Research Challenges for HumanComputer Interaction. *International Journal of Human-Computer Interaction 32*(8), 585–612.

Preis, T., H. S. Moat, and H. Eugene Stanley (2013). Quantifying trading behavior in financial markets using google trends. *Scientific Reports 3*.

Prensky, M. (2003, oct). Digital game-based learning. *Computers in Entertainment 1*(1), 21.

Prestopnik, N. and K. Crowston (2012). Purposeful Gaming & Socio-Computational Systems : A Citizen Science Design Case. *Group'12*, 75–84.

Prestopnik, N., K. Crowston, and J. Wang (2014). Exploring Data Quality in Games With a Purpose. *iConference 2014 Proceedings* (2008), 1–15.

Pretty, J. (2003). Social Capital and the Collective Management of Resources. *Science 302*(5652), 1912–1914.

Price, C. A. and H. S. Lee (2013). Changes in participants' scientific attitudes and epistemological beliefs during an astronomical citizen science project. *Journal of Research in Science Teaching 50*(7), 773–801.

Price, D. J. d. S. and D. d. B. Beaver (1966). Collaboration in an Invisible College. *American Psychologist 21*(11), 1011–1018.

Pulcu, E., E. J. Thomas, P. D. Trotter, M. McFarquhar, G. Juhasz, B. J. Sahakian, J. F. W. Deakin, I. M. Anderson, R. Zahn, and R. Elliott (2014). Social-economical decision making in current and remitted major depression. *Psychological Medicine 45*(6), 1–13.

Qaurooni, D., A. Ghazinejad, I. Kouper, and H. Ekbia (2016). Citizens for Science and Science for Citizens: The View from Participatory Design. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (December 2017), 1822–1826.

Raddick, M. J., G. Bracey, K. Carney, and G. Gyuk (2010). Citizen science: status and research directions for the coming decade. *AGB Stars and Related Phenomonastro 2010: The Astronomy and Astrophysics Decadal Survey*, 46.

Raddick, M. J., G. Bracey, P. L. Gay, C. J. Lintott, P. Murray, K. Schawinski, A. S. Szalay, and J. Vandenberg (2010). Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Astronomy Education Review 9*(1), 010103.

Raddick, M. J., J. Hopkins, G. Bracey, and P. L. Gay (1988). Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. (Root).

Radford, J., A. Pilny, A. Reichelmann, B. Keegan, B. F. Welles, J. Hoye, K. Ognyanova, W. Meleis, and D. Lazer (2016). Volunteer Science. *Social Psychology Quarterly 79*(4), 376–396.

Ralambondrainy, H. and H. (1995, nov). A conceptual version of the K-means algorithm. *Pattern Recognition Letters 16*(11), 1147–1157.

Rand, D. G., J. D. Greene, and M. A. Nowak (2012). Spontaneous giving and calculated greed. *Nature 489*(7416), 427–430.

Rand, D. G., A. Peysakhovich, G. T. Kraft-Todd, G. E. Newman, O. Wurzbacher, M. A. Nowak, and J. D. Greene (2014). Social heuristics shape intuitive cooperation. *Nature Communications 5*, 1–12.

Rapoport, A. and A. M. Chammah (1965). *Prisoner's Dilemma*. The University of Michigan Press.

Rapoport, A. and A. M. Chammah (1966). The Game of Chicken. *American Behavioral Scientist 10*(3), 10–28.

Rapoport, A. and M. Guyer (1966). *A Taxonomy of 2 x 2 Games*. Bobbs-Merrill Reprint Series in the Social Sciences.

Raupach, M. R., S. J. Davis, G. P. Peters, R. M. Andrew, J. G. Canadell, P. Ciais, P. Friedlingstein, F. Jotzo, D. P. van Vuuren, and C. Le Quéré (2014). Sharing a quota on cumulative carbon emissions. *Nature Climate Change 4*(10), 873–879.

Rebecca Boyle (2016). A.I. Is Getting Better at Spotting Galaxies. *The Atlantic*.

Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology 49*(4), 243–256.

Rhodes, C. J. (2016, mar). The 2015 Paris Climate Change Conference: COP21. *Science Progress 99*(1), 97–104.

Rilling, J. K., A. L. Glenn, M. R. Jairam, G. Pagnoni, D. R. Goldsmith, H. A. Elfenbein, and S. O. Lilienfeld (2007). Neural Correlates of Social Cooperation and Non-Cooperation as a Function of Psychopathy. *Biological Psychiatry 61*(11), 1260–1271.

Rosenthal, R. and R. L. Rosnow (1969). *Artifact in behavioral research*. New York: Academic Press.

Rotman, D., J. Preece, J. Hammock, K. Procita, D. Hansen, C. Parr, D. Lewis, and D. Jacobs (2012). Dynamic Changes in Motivation in Collaborative Citizen-Science Projects. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, 217–226.

Roy, H. E., M. J. O. Pocock, C. D. Preston, D. B. Roy, J. Savage, J. C. Tweddle, L. D. Robinson, Roy, H.E., M.J.O. Pocock, C.D. Preston, J. Savage, J.C. Tweddle, and L. Robinson (2012). Understanding Citizen Science and Environmental Monitoring. *UK Environmental Observation Framework*, 170.

Rudin, C. (2013). Predictive policing: Using machine learning to detect patterns of crime. *Wired*, 1–5.

Rusbult, C. E. and P. A. M. Van Lange (2003, feb). Interdependence, Interaction, and Relationships. *Annual Review of Psychology 54*(1), 351–375.

Ruttan, L. (2006). Sociocultural Heterogeneity and the Commons. *Current Anthropology 47*(5), 843–853.

Ryan, R. M. and E. L. Deci (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*.

Sagarra, O., M. Gutiérrez-Roig, I. Bonhoure, and J. Perelló (2016, jan). Citizen Science Practices for Computational Social Science Research: The Conceptualization of Pop-Up Experiments. *Frontiers in Physics 3*(January), 1–19.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly.

Sánchez, A. (2018, feb). Physics of human cooperation: experimental evidence and theoretical models. *Journal of Statistical Mechanics: Theory and Experiment 2018*(2), 024001.

Santos, F. C., V. V. Vasconcelos, M. D. Santos, P. N. B. Neves, and J. M. Pacheco (2012). Evolutionary dynamics of climate change under collective-risk dilemmas. *Mathematical Models and Methods in Applied Sciences 22*(supp01), 1140004.

Sauermann, H. and C. Franzoni (2015). Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences*, 201408907.

Scheele, D., Y. Mihov, O. Schwederski, W. Maier, and R. Hurlemann (2013). A negative emotional and economic judgment bias in major depression. *European Archives of Psychiatry and Clinical Neuroscience 263*(8), 675–683.

SciStarter (2017). SciStarter.

Senabre, E., N. Ferran-Ferrer, and J. Perelló (2018). Participatory design of citizen science experiments. *Comunicar 26*(54), 29–38.

Shannon, C. E. (1950). Programming a computer for playing chess. *Philosophical Magazine 41*(314), 256.

Shao, R., H.-j. Zhang, and T. M. C. Lee (2015). The neural basis of social risky decision making in females with major depressive disorder. *Neuropsychologia 67*, 100–10.

Shapiro, R. B. and K. D. Squire (2011). Games for Participatory Science: A Paradigm for Game-based Learning for Promoting Science Literacy.

Sheppard, S. A. and L. Terveen (2011). Quality is a Verb : The operationalization of data quality in a citizen science community. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym '11)*, 29–38.

Shiller, R. (1992). *Market volatility.* MIT Press.

Shirk, J. L., H. L. Ballard, C. C. Wilderman, T. Phillips, A. Wiggins, R. Jordan, E. McCallie, M. Minarchek, B. V. Lewenstein, M. E. Krasny, and R. Bonney (2012). Public participation in scientific research: A framework for deliberate design. *Ecology and Society 17*(2).

Silver, D., T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis (2017). Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. pp. 1–19.

Silvertown, J. (2009). A new dawn for citizen science. *Trends in ecology & evolution 24*(9), 467–71.

Simon Rogers (2013). John Snow's data journalism: the cholera map that changed the world.

Snow, C. (1959). *The Two Cultures and the Scientific Revolution.* Cambridge University Press.

Soumith Chintala and Yann LeCun (2016). A path to unsupervised learning through adversarial networks.

Stagnaro, M. N., A. A. Arechar, and D. G. Rand (2017, oct). From good institutions to generous citizens: Top-down incentives to cooperate promote subsequent prosociality but not norm enforcement. *Cognition 167*, 212–254.

Steel, Z., C. Marnane, C. Iranpour, T. Chey, J. W. Jackson, V. Patel, and D. Silove (2014). The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology 43*(2), 476–493.

Stella, M., E. Ferrara, and M. De Domenico (2018, feb). Bots sustain and inflate striking opposition in online social systems. *arXiv*.

Stoop, J. (2014). From the lab to the field: Envelopes, dictators and manners. *Experimental Economics 17*(2), 304–313.

Sugden, R. (2004). *The economics of rights, co-operation and welfare.*

Suri, S. and D. J. Watts (2011, mar). Cooperation and contagion in web-based, networked public goods experiments. *PLoS ONE 6*(3), e16836.

Surowiecki, J. (2004). The Wisdom of Crowds: Why the Many are Smarter Than the Few and How ... *How Collective Wisdom Shapes Business Economies Societies and Nations New York Doubleday 27*(5), 296.

Tavoni, A., A. Dannenberg, G. Kallis, and A. Löschel (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *PNAS 108*(29), 11825–11829.

Thaler, R. H. (2000). From Homo Economicus to Homo Sapiens. *Journal of Economic Perspectives 14*(1), 133–141.

The Behavioural Insights Team (2018). The Behavioural Insights Team.

The White House (2009). Memorandum on Transparency and Open Government.

The White House (2013). Memorandum on Open Data PolicyManaging Information as an Asset.

Tibshirani, R., G. Walther, and T. Hastie (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(2), 411–423.

Tinati, R., M. Luczak-Roesch, E. Simperl, and W. Hall (2017, jan). An investigation of player motivations in Eyewire, a gamified citizen science project. *Computers in Human Behavior*.

Trumbull, D. J., R. Bonney, D. Bascom, and A. Cabral (2000, mar). Thinking scientifically during participation in a citizen-science project. *Science Education 84*(2), 265–275.

Trump, D. (2017). Statement by President Trump on the Paris Climate Accord.

Tversky, A. and D. Kahneman (1974, sep). Judgment under Uncertainty: Heuristics and Biases. *Science 185*(4157), 1124–1131.

Tweddle, J., L. Robinson, M. Pocock, and H. Roy (2012). *Guide to citizen science: developing, implementing and evaluating citizen science to study biodiversity and the environment in the UK*. Number 10. Natural History Museum and NERC Centre for Ecology & Hydrology for UK-EOF.

USGS (2017). American Breeding Bird Survey.

Van Lange, P. A. M. (2000, jan). Beyond Self-interest: A Set of Propositions Relevant to Interpersonal Orientations. *European Review of Social Psychology 11*(1), 297–331.

Van Lange, P. A. M., J. Joireman, C. D. Parks, and E. Van Dijk (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes 120*(2), 125–141.

Vasconcelos, V. V., F. C. Santos, J. M. Pacheco, and S. A. Levin (2014). Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences 111*(6), 2212–2216.

Verhaak, R. G., K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B. A. Weir, S. Gabriel, W. Winckler, S. Gupta, L. Jakkula, H. S. Feiler, J. G. Hodgson, C. D. James, J. N. Sarkaria, C. Brennan, A. Kahn, P. T. Spellman, R. K. Wilson, T. P. Speed, J. W. Gray, M. Meyerson, G. Getz, C. M. Perou, and D. N. Hayes (2010, jan). Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell 17*(1), 98–110.

Vicens, J., N. Bueno-Guerra, M. Gutiérrez-Roig, C. Gracia-Lázaro, J. Gómez-Gardeñes, J. Perelló, A. Sánchez, Y. Moreno, and J. Duch (2017). Resource heterogeneity leads to unjust effort distribution in climate change mitigation. *arXiv (manuscript submitted for publication).*

Vicens, J., J. Perelló, and J. Duch (2018). Behavioural experimental platform. *Manuscript submitted for publication.*

Vicens, J., H. Zhang, and J. Duch (2018). Natural Patterns: A Platform To Enhance Participation And Science Disposition In Citizen Science Activities. *Manuscript submitted for publication.*

Villafuerte, L., M. Markova, and S. Jorda (2012). Acquisition of social abilities through musical tangible user interface. *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*, 745.

Vitone, T., K. A. Stofer, M. S. Steininger, J. Hulcr, R. Dunn, and A. Lucky (2016). School of ants goes to college: Integrating citizen science into the general education classroom increases engagement with science. *Journal of Science Communication 15*(1), 1–24.

Von Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. *Princeton University Press*, 625.

Wang, Y., L.-Q. Yang, S. Li, and Y. Zhou (2015). Game Theory Paradigm: A New Tool for Investigating Social Dysfunction in Major Depressive Disorders. *Frontiers in Psychiatry 6*(September), 128.

Wang, Z., B. Xu, and H. J. Zhou (2014). Social cycling and conditional responses in the Rock-Paper-Scissors game. *Scientific Reports 4*.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association 58*(301), 236.

Watercutter, A. (2017). You're Never Too Old to Color - Especially Math Patterns.

Wattanasoontorn, V., I. Boada, R. García, and M. Sbert (2013, dec). Serious games for health. *Entertainment Computing 4*(4), 231–247.

Watts, D. J. (2007, feb). A twenty-first century science.

Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour 1*(1).

West, S. and R. Pateman (2016, dec). Recruiting and Retaining Participants in Citizen Science: What Can Be Learned from the Volunteering Literature? *Citizen Science: Theory and Practice 1*(2).

White, R. G. E., S. E. Jain, D. M. R. E. Orr, and U. M. E. Read (2017). The Palgrave handbook of sociocultural perspectives on global mental health.

Wiggins, A. and K. Crowston (2011). Goals and tasks: Two typologies of citizen science projects. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3426–3435.

Wilkerson, M. D. and D. N. Hayes (2010, jun). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics 26*(12), 1572–1573.

Wolch, J. R., J. Byrne, and J. P. Newell (2014, may). Urban green space, public health, and environmental justice: The challenge of making cities just green enough'. *Landscape and Urban Planning 125*, 234–244.

World Health Organisation (2001). The World Health Report 2001: mental health, new understanding, new hope.

Xavier Amatriain and Justin Basilico (2012). Netflix Recommendations: Beyond the 5 stars.

Young, H. P. (2015, aug). The Evolution of Social Norms. *Annual Review of Economics 7*(1), 359–387.

Zhou, Y., H. Yu, and X. Cai (2009, dec). A Novel k-Means Algorithm for Clustering and Outlier Detection. In *2009 Second International Conference on Future Information Technology and Management Engineering*, pp. 476–480. IEEE.

enbabaolu, Y., G. Michailidis, and J. Z. Li (2015, may). Critical limitations of consensus clustering in class discovery. *Scientific Reports 4*(1), 6207.

Appendices

Appendix A

Mr.Banks and Dr.Brain

A.1 Mr.Banks

A.1.1 Data repository

The data collected in the experiment is available in the public repository (Gutiérrez-Roig et al., 2016a): Dataset Mr.Banks[1].

A.1.2 Tutorial

Screenshots of the tutorial shown before the experiment starts: Tutorial Mr.Banks[2].

A.2 Dr.Brain

A.2.1 Data repository

The data collected in the experiment is available in the public repository (Poncela-Casasnovas et al., 2017): Dataset Dr.Brain[3]. .

A.2.2 Tutorial

Screenshots of the tutorial shown before the experiment starts: Tutorial Dr.Brain[4].

A.2.3 Translated transcript of the tutorial

Here, we present the translation into English of the tutorial text (the original was made available to the participants in Castilian/Spanish and Catalan.

Tutorial Screen #1.

> *Welcome to Dr. Brain. The game, designed to study how we make decisions, is made of several rounds with different opponents located in the DAU. During the experiment we don't expect you to behave in any particular way: there are no wrong nor incorrect answers. You will simply have a limited time to*

1. https://doi.org/10.5281/zenodo.50429
2. https://github.com/jvicens/Thesis/tree/master/Appendix/Screenshots
3. https://doi.org/10.5281/zenodo.1127154
4. https://github.com/jvicens/Thesis/tree/master/Appendix/Screenshots

make your decisions. In these next screens we will teach you how to play Dr. Brain. Use the side arrow keys to move within the tutorial, and when you are done you will be able to start the rounds. This game has been thought by scientists from the Universitat of Barcelona (UB), Universitat Rovira i Virgili (URV), Instituto de Biocomputación and Sistemas Complejos (BIFI)-Universidad de Zaragoza (UZ) and Universidad Carlos III in Madrid (UC3M). It is an experiment to study and understand how we humans make decisions.

Tutorial Screen #2.

The rules of Dr. Brain. It is important that you don't talk to other players during the experiment. Keep focused! The decisions made during the experiment and the accumulated points will determine your chances of wining prizes: the more points, the more tickets you will get for the raffle. If you leave the game while it is in progress, you won't be able to come back in!

Tutorial Screen #3.

This is the screen you will see when the rounds of the game start. In each one of them, we will assign you a random partner to play.

Tutorial Screen #4..

Each round has a table that represents your opponent's possible actions as well as yours. Your opponent and you will follow the same rules in the round. In this way, depending on what each one of you choose, you will win more or less. The rows represent your choice, the columns represent your opponent's. For each choice, it is listed how much you will win, and how much your opponent will.

Tutorial Screen #5.

Pay attention, the tables may change from round to round, and the rules may be different. You may win more or less points, o what seemed more interesting may be different now.

Tutorial Screen #6.

To play you must choose one of the two options, represented by a color. Your opponent plays following the same rules as you, described in the table, but you won't know his choice until after the end of the round.

Tutorial Screen #7.

Every round of the game lasts 40 seconds, you have to choose one of the two actions during that time. If you don't choose anything, the computer will do it for you randomly and you will move on to play the next round. Don't worry, 40 seconds is plenty of time!

Tutorial Screen #8.

Example: If you pick RED and your opponent picks GREEN. You (red) win 8 and your opponent (green) wins 6.

Tutorial Screen #9.

Example: If you pick PURPLE and your opponent picks YELLOW. You (purple) win 11 and your opponent (yellow) wins 0. If your adversary chooses... If you choose... You win... He wins... What do you choose?

Appendix B

The Climate Game

B.1 Data repository

The data collected in the experiment will be available once the scientific publication is effective.

B.2 Tutorial

Screenshots of the tutorial shown before the experiment starts: Tutorial Dr.Brain: The Climate Game[1].

B.3 Transcript of the tutorial

Before the collective-risk dilemma started, we showed to each group of subjects a tutorial that was included in the same platform used to participate in the game. The subjects were assisted by researchers who answered any questions that came up about the experiment. The tutorial was presented identically in three languages (Catalan, Spanish and English). Here we show the tutorial screens and text of the English version.

Tutorial Screen #1.

> *Welcome to The Climate Game. TUTORIAL: HOW DO YOU PLAY?. This screen will show you how to play the game suggested by Dr.Brain, a game designed to study how we make decisions. This game is designed by scientists from the Universitat de Barcelona (UB), Universitat Rovira i Virgili (URV), Instituto de Biocomputación y Sistemas Complejos (BIFI), Universidad de Zaragoza (UZ) and Universidad Carlos III de Madrid (UC3M), is an experiment to study and understand how human make decisions. Use the lateral arrows to side-scroll and navigate in the tutorial, once you finish you can start the game.*

Tutorial Screen #2.

> *It is important that you do not talk with the other players during the experiment. 2) We do not expect you to behave in any special way: there are no right or wrong answers. 3) If you exit the game while the game is running, you can not re-enter! 4) The decisions taken during the game will have real consequences in both the money that you get at the end of the game and in the financing of actions against climate change.*

Tutorial Screen #3.

1. https://github.com/jvicens/Thesis/tree/master/Appendix/Screenshots

In the game you will play with 5 other people to be randomly selected so that no one will know who you are playing with

Tutorial Screen #4.

Before the game starts, Dr. Brain will randomly assign you a player number and the amount of money that you will have initially. This initial capital will be from 20 to 60 euros. Also, you will know the initial capital of the other players!

Tutorial Screen #5.

The target of the game is to raise 120 euros in a common fund to finance actions against climate change. The game will run for 10 rounds. In each round each player has to contribute from 0 to 4 euros of their own capital to the common fund.

Tutorial Screen #6.

At the end of each round and once the six players have decided, you will see: 1) The amount of money that is in the common fund. 2) How much has each player contributed in the round. 3) The starting and current capital of each player.

Tutorial Screen #7.

In each round you have 30 seconds to make a decision. If after this time you have not decided, the computer will do it in your place. Important! If time runs out in two or more rounds you will not get any profit. Stay focused!

Tutorial Screen #8.

If after 10 rounds THERE ARE 120 EUROS OR MORE in the common fund: 1. The participants will recieve a gift card amounting to the value of their savings. 2. We will fund actions against climate change, such as the reforestation of "Parc de Collserola" in Barcelona. If after 10 rounds THERE ARE LESS THAN 120 EUROS in the common fund: 1. There is a 10% possibility that the participants will receive their savings as a gift card. 2. We will not be able to spend money on climate change actions, such as planting trees.

Tutorial Screen #9.

The Climate Game. Once you finish the game, you will see a screen with the final result and then we will ask you to fill out a short survey. Remember, the outcome depends on your decisions and those of the rest of your peers. If you have any questions you can ask any of the organizers. Touch the button to continue

B.4 Questionnaire participation and climate change

After the collective-risk dilemma ended, as a last stage of the experiment, for each group of subjects we asked a questionnaire with two sets of questions: (i) how they made decisions in the game and (ii) some basic concepts about climate change. The questionnaire was also presented identically in three languages (Catalan, Spanish and English), here we present the questions in English.

Question #1

Had you previously participated in Citizen Science experiments?

a.*No*; b.*Yes*; c.*Yes, in previous DAU experiments.*

Question #2:

Did you like the experience?

a.*Very much*; b.*Somewhat*; c.*Not really*; d.*Not at all.*

Question #3:

At the beginning of the experiment, did you expect to reach the 120€ target?

a.*Yes, from the beginning*; b.*Yes, after a few rounds*; c.*No, from the beginning*; d.*No, after a few rounds.*

Question #4:

Generally, if others are contributing little, I should also contribute little.

a.*Agree*; b.*Disagree*; c.*My contribution should not depend on this*; d.*n/a.*

Question #5:

Generally, if others are contributing a lot, I should also contribute a lot.

a.*Agree*; b.*Disagree*; c.*My contribution should not depend on this*; d.*n/a.*

Question #6:

I think there have been players who have taken advantage of the generosity of others to maintain their capital.

a.*Agree*; b.*Disagree*; c.*n/a.*

Question #7:

The contributions of each player should be proportional to their capital: those who have started with more should contribute more and those who have started with less should contribute less.

a.*Agree*; b.*Disagree*; c.*n/a.*

Question #8:

The contributions of each player should be fair, because the benefits of the common fund affect everyone equally, as well as the risk of losing everything.

a.*Agree*; b.*Disagree*; c.*n/a.*

Question #9:

It seems fair to me that a player with a lot of capital should get money at the end, if he has contributed at least half of what he had.

a.*Agree*; b.*Disagree*; c.*n/a.*

Question #10:

When the polluting gases prevent the rays of the Sun from coming out of the Earth, it is due to...

a.*Doppler Effect*; b.*Greenhouse Effect*; c.*Faraday Effect*; d.*Refrigerator Effect.*

Question #11:

Which of the following countries is the most polluting in the world?

a.*U.S*; b.*Italy*; c.*China*; d.*Japan.*

Question #12:

Which of the following elements is the least polluting?

a.*Oil*; b.*Carbon*; c.*Solar energy*; d.*Nuclear energy.*

Question #13:

Which international treaty tries to regulate CO2 emissions to the atmosphere?

a.*Declaration of Helsinki*; b.*Kyoto Protocol*; c.*Schengen Agreement*; d.*Treaty of Versailles.*

Question #14:

What is the total number of gaseous pollutants emitted by each individual?

a.*Carbon footprint*; b.*Eco-Impact*; c.*Individual gas fee*; d.*Reduced environmental cost.*

Question #15:

According to the economist Nicolas Stern, if urgent measures are not implemented, what costs could climate change represent in 2050, as a percentage of world GDP?

a.*2%*; b.*5%*; c.*15%*; d.*20%*

Appendix C

Detailed Results in the Fight Against Climate Change

C.1 Payoff

Table C.1: Payoff in Climate Game. Payoff and payoff normalized by relative fairness

			Payoff			Payoff Normalized		
Endowment	**Treatment**	**n**	**Mean**	**SD**	**SE**	**Mean**	**SD**	**SE**
20	Unequal	26	5.92	5.56	1.09	0.59	0.56	0.11
30	Unequal	27	9.93	7.75	1.49	0.66	0.52	0.10
40	Unequal	54	18.57	7.45	1.01	0.93	0.37	0.05
40	Equal	159	18.34	5.8	0.46	0.92	0.29	0.02
50	Unequal	27	23.74	6.93	1.33	0.95	0.28	0.05
60	Unequal	27	32	6.11	1.18	1.07	0.2	0.04

Table C.2: Payoff differences in Climate Game. Pairwise comparation of payoff normalized by relative fairness.

Endowment	**20‡**	**30‡**	**40†**	**40‡**	**50‡**	**60‡**
20‡	-	**-0.5 (0.640)**	-2.9** (0.007)	-2.8** (0.008)	-2.9** (0.005)	-4.1*** ($3 \cdot 10^{-4}$)
30‡	**0.5 (0.640)**	-	-2.5* (0.018)	-2.4* (0.021)	-2.5* (0.015)	-3.8** (0.001)
40†	2.9** (0.007)	2.5* (0.018)	-	**-0.2 (0.834)**	**-0.6 (0.578)**	-3.3** (0.002)
40‡	2.8** (0.008)	2.4* (0.021)	**0.2 (0.834)**	-	**-0.3 (0.777)**	-2.1* (0.034)
50‡	2.9** (0.005)	2.5* (0.015)	**0.6 (0.578)**	**0.3 (0.777)**	-	**-1.8 (0.083)**
60‡	4.1*** ($3 \cdot 10^{-4}$)	3.8** (0.001)	3.3** (0.002)	2.1* (0.034)	**1.8 (0.083)**	-

† Equal treatment and ‡ Unequal treatment. The table displays the values of Welch t-tests (p-value) used to compare the differences of payoff's means among the groups. *significant at 5%; **significant at 1%; *** significance less than 0.1%.

C.2 Cohort analysis effect of minors.

Table C.3: Cohort analysis of game contributions in games with and without minors. There are no significant differences in the game contributions ($p > 0.05$).

Treatment	Minor	Mean	SD	n	SE	t	df	p-value
Both	Yes	129.7	10.2	29	1.9	0.8	44.6	0.45
	No	132.2	13.4	25	2.7			
Equal	Yes	127.1	5.7	15	1.5	1.3	13.2	0.22
	No	133.4	16.1	12	4.6			
Unequal	Yes	132.4	13.3	14	3.5	-0.3	24.6	0.776
	No	131.0	10.9	13	3.0			

Table C.4: Cohort analysis of contribution per endowment in groups of minors and adults. There are no significant differences in the contributions per endowment ($p > 0.05$).

Endowment	Treatment	Minor	Mean	SD	n	SE	t	df	p-value
20	Unequal	Yes	17	2.6	3	1.5	-1.7	5	0.149
		No	13.7	5.8	23	1.2			
30	Unequal	Yes	23.2	7.3	4	3.6	-0.9	4.3	0.399
		No	19.5	7.8	23	1.6			
40	Unequal	Yes	22.6	7	10	2.2	-0.6	14.1	0.575
		No	21.1	7.6	44	1.1			
40	Equal	Yes	21.4	6.2	28	1.2	0.2	37.4	0.801
		No	21.7	5.7	131	0.5			
50	Unequal	Yes	27.2	9.7	4	4.8	-0.2	3.5	0.830
		No	26.1	6.6	23	1.4			
60	Unequal	Yes	25.5	3.9	4	1.9	1.2	6.3	0.256
		No	28.4	6.4	23	1.3			

Appendix D

Games for Mental Health

D.1 Data repository

The data collected in the experiment is available in the public repository (Cigarini et al., 2018a): Dataset Games for Mental Health[1].

D.2 Tutorial

Screenshots of the tutorial shown before the experiment starts: Tutorial Games for Mental Health[2].

D.3 Translated transcript of the tutorial

Tutorial Screen #1.

> ***Tutorial: How to play***
>
> *The activity you are about to engage consists of a set of games that you will be presented. It is extremely important that during the experiment you DO NOT TALK with other players. You are not expected to behave in any particular way: there are no right or wrong answers. If you exit the game before its end, you won't be able to enter again! The money you will get at the end of the game will depend on the decisions you will make throughout the experiment. *Use the side arrows to move through the tutorial. Once you are done, you will start with the first game.*

Tutorial Screen #2-7.

> ***Game One: The Climate Game***
>
> *In this game you will simultaneously play with 5 PLAYERS. Each player will be endowed with 40 MONETARY UNITS. The objective is to collect 120 MONETARY UNITS in a common fund to promote actions against climate change.*
>
> *The game consists of 10 rounds. In each round every player has to contribute between 0 and 4 monetary units of her actual endowment to the common fund.*

1. https://doi.org/10.5281/zenodo.1175627
2. https://github.com/jvicens/Thesis/tree/master/Appendix/Screenshots

You have 20 seconds to make your decision, or else the computer will do it for you. If you run out of time for 2 or more rounds your earnings will be 0 in this game.

At the end of each round you will be presented with a summary containing information on: 1. The total amount of money collected in the common fund. 2. The contribution of every player in the previous round. 3. The initial and current endowment of every player.

If, after 10 rounds THERE ARE 120 OR MORE MONETARY UNITS in the common fund: THE MONETARY UNITS YOU SAVED will add on a voucher for ABACUS shop. We will support actions to promote reforestation in Catalunya.

Example: The game ends and the whole group contributed 130 monetary units. If you contributed with 26 monetary units out of your initial 40, you will earn a voucher with the equivalent of the remaining 14 monetary units!

If, after 10 rounds THERE ARE LESS THAN 120 MONETARY UNITS in the common fund: 1.With a 10% probability you will earn a voucher for ABACUS shop with the equivalent of the MONETARY UNITS YOU SAVED. 2.We will not be able to promote actions for reforestation.

Example: The game ends and the whole group contributed 94 monetary units. If you contributed with 12 monetary units out of your initial 40, you will earn a voucher with the equivalent of your remaining 28 monetary units only with a 10% probability. In 90% of cases you will not earn anything.

Tutorial Screen #8-11.

Game Two: Investor Game

You are faced with a good dilemma. You can be either an investor who has to decide whether to invest in a new business. Or you can happen to be an entrepreneur who needs money for his enterprise. What role will you play? What are you going to do? Think carefully about which decision to take in every situation.

You are: The investor! Your game partner is an entrepreneur asking you to invest in his enterprise. With the money you give him he will start a business which will earn him THREE TIMES your investment. In this game you earn the money you will not invest plus the money that the entrepreneur will return you. He will have to decide which portion of the money earned to return you, but has NO OBLIGATION to reciprocate. How much money do you want to invest in the business?

[Selection]

You are: The entrepreneur! Your game partner is an investor with 10 MONEY UNITS, and has to decide how much money to invest in your business. The INVESTOR decided to invest in your business: ... Thanks to the investment you earned: ... Which portion of the money earned do you want to return him?

[Selection]

Tutorial Screen #12-15.

Game Three: Prize Game

You and the other player received a prize worth 10 MONEY UNITS each. Now you both need to make a decision which will affect what you both will earn: KEEP the prize or MULTIPLY it, taking part of the prize of the other player. It is not that easy at it seems. According to what the other player does you might not be earning anything.

The rules are as follows:
If you choose KEEP and the other chooses KEEP:
You earn 10 and he/she earns 10.
If you choose KEEP and the other chooses MULTIPLY:
You earn 5 and he/she earns 15.
If you choose MULTIPLY and the other chooses KEEP:
You earn 15 and he/she earns 5.
If you choose MULTIPLY and the other chooses MULTIPLY:
You earn 0 and he/she earns 0.

Before playing the game... A CONFESSION: What do you think the other player will do?

[Selection]

Now it's your turn... What do you want to do with your prize?

[Selection]

D.4 Questionnaire sociodemographic

Question #1

Why are you being looked after?
a.*Depression* b.*Bipolar disorder* c.*Psychosis (schizophrenia, etc.)* d.*Anxiety* e.*Other mental disorders* f.*I do not have any mental disorder*

Question #2 (if "I do not have any mental disorder")

Are you?
a.*Non-caregiving relative* b.*Caregiving relative or informal caregiver* c.*Friend* d.*Professional (health sector, social sector, etc.)* e.*Others*

Question #3

Employment status
a.*Not working* b.*Working* c.*Working in CET (Centro Especial de Trabajo)*

Question #4

Gender

a.*Man* b.*Woman*

Question #5

Enter your age

Question #6

Enter your Postal Code

Question #7

Civil status
a.*Single* b.*Married* c.*Civil union* d.*Divorced* e.*Widow* f.*Others*

Question #8

Educational level
a.*None* b.*Primary education* c.*Secondary education* d.*Baccalaureate or equivalent* e.*Professional education* f.*University studies* g.*Not specified*