Quantitative account of social interactions in a mental health care ecosystem: cooperation, trust and collective action

Anna Cigarini[1,2], Julián Vicens[1,2,3], Jordi Duch[3,4,8], Angel Sánchez[5,6,7], and Josep Perelló[1,2,*]

[1]Departament de Física de la Matèria Condensada, Universitat de Barcelona, 08028 Barcelona, Spain
[2]Universitat de Barcelona Institute of Complex Systems UBICS, 08028 Barcelona, Spain
[3]Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007, Tarragona, Spain
[4]Grupo Interdisciplinar de Sistemas Complejos (GISC), Unidad de Matemática, Modelización y Ciencia Computacional, Universidad Carlos III de Madrid, 28911 Leganés, Spain
[5]Unidad Mixta Interdisciplinar de Comportamiento y Complejidad Social (UMICCS) UC3M-UV-UZ, Universidad Carlos III de Madrid, 28911 Leganés, Spain
[6]Institute UC3M-BS of Financial Big Data, Universidad Carlos III de Madrid, 28903 Getafe, Spain
[7]Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, 50009 Zaragoza, Spain
[8]Northwestern Institute on Complex Systems (NICO), Northwestern University, 60208 Evanston, IL, USA
[*]josep.perello@ub.edu

ABSTRACT

Mental disorders have an enormous impact in our society, both in personal terms and in the economic costs associated with their treatment. In order to scale up services and bring down costs, administrations are starting to promote social interactions as key to care provision. We analyze quantitatively the importance of communities for effective mental health care, considering the involvement of all community members involved. 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. The evidence reinforces the idea of community social capital, with caregivers and professionals playing a leading role. Yet, the cost of collective action is mainly supported by individuals with a mental condition - which unveils their vulnerability. The results are in general agreement with previous findings but, since we broaden the perspective of previous studies, we are also able to find marked differences in the social behavior of certain groups of mental disorders. We finally point to the conditions under which cooperation among members of the ecosystem is better sustained, suggesting how virtuous cycles of inclusion and participation can be promoted in a 'care in the community' framework.

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[1]. 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[2]. As part of a global effort to scale up services and bring down costs, reliance is increasingly made upon informal social networks[3]. 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[4]. 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[5]: 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[6–8]. 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[9,10]. Most research addressed the issue of behavioral differences between individuals with MD and healthy populations[6,7,11–16]. These works, that point to cognitive and affective processing impairments[6,16,17], 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 primarly 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.

We here adopt a novel community perspective. Our objective is twofold: First, we aim to develop a thorough taxonomy of the behavioral traits of role groups within the collective. We thus account for both the heterogeneity of actors, and for multiple types of social interactions. We strongly believe that to predict and understand behavior is necessary to consider the relationship context in which individuals are embedded. Therefore diversity of roles, motivations or capabilities, must be taken into account. Also, real life social interactions occur in different forms; sometimes people must work together, some others they have to coordinate or anti-coordinate their behaviors, yet in other situations they find themselves in more or less disadvantaged positions. It is therefore of crucial importance to encompass a comprehensive range of strategic situations if we are to appreciate behavior. That is, traits such as trust, altruism, or reciprocity, along with the person's own expectations, all play a role in the process of decision making in social contexts. This calls for an experimental approach in which participants face several strategic settings. Our second objective is to provide quantitative accounts of social capital within the mental health community, bringing the notion of social capital into the forefront of mental health care. Far from being universally defined, its core contention is that social networks are a valuable asset, providing a basis for social cohesion and cooperation towards a common goal[18] (which is, in our case, mental care provision). It thus encompasses those norms and forces that shape social interactions, serving as the glue that holds society together[19].

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[20], in order to be as close as possible to the dynamic and unique nature of real-life social interactions. The design of our socialized setup is based on a participatory process which counted on the collaboration of all stakeholders of the mental health ecosystem. By combining all these ingredients, we have developed a framework that, as will be shown below, 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.

Results

We begin the presentation of our results from the dyadic games of our suite of strategic interactions. Aggregate behavioral measures point to systematic deviations from self-interested predictions which are in line with previous literature on experimental game play[21]. In the Prisoner's Dilemma (PD), the average cooperation rate across all individuals is $c = 0.61 \pm 0.03$ (standard error of the mean), which is notoriously well above the Nash equilibrium prediction of $c = 0$. Participants behavior in the PD is also significantly associated with their estimates about the likely cooperation of the partner ($\chi^2 = 32.48$, $p = 1.2 \cdot 10^8$), with 44% of all participants expecting the partner to cooperate, and thus cooperating themselves. This points to the crucial role of positive expectations on cooperative behavior[22]. Further, participants trust and reciprocate positive amounts in the Trust Game (5.79 ± 0.15 monetary units (MU) and $41.3 \pm 1.37\%$ of the amount available to return, respectively), again departing largely from Nash equilibrium conjectures of 0 MU transferred. The results also suggest that in considering the mental health community in its whole, thus accounting for the diversity of actors and roles, the global picture does not substantially differ from society at large.

Sectorial and dyadic behavior

As we stated above, our main interest is to delve into the behavior of the different actors who make up the mental health ecosystem Figure 1 summarizes the results for the five groups of individuals concerned. The heatmap yields several insights that are worth commenting upon.

In one-shot dyadic interactions some marked differences in the frequency of cooperative behaviors (PD) arise within the collective formed by affected with MD, caregivers, non-caregivers (Kruskal-Wallis rank sum test, $H = 6.04, df = 2, p = 0.0488$). Further pairwise comparisons (see Supplementary Table 1) show that participants with anxiety and caregivers are more likely to opt for the cooperative strategy compared to participants with bipolar disorder, psychosis or other members of the collective.

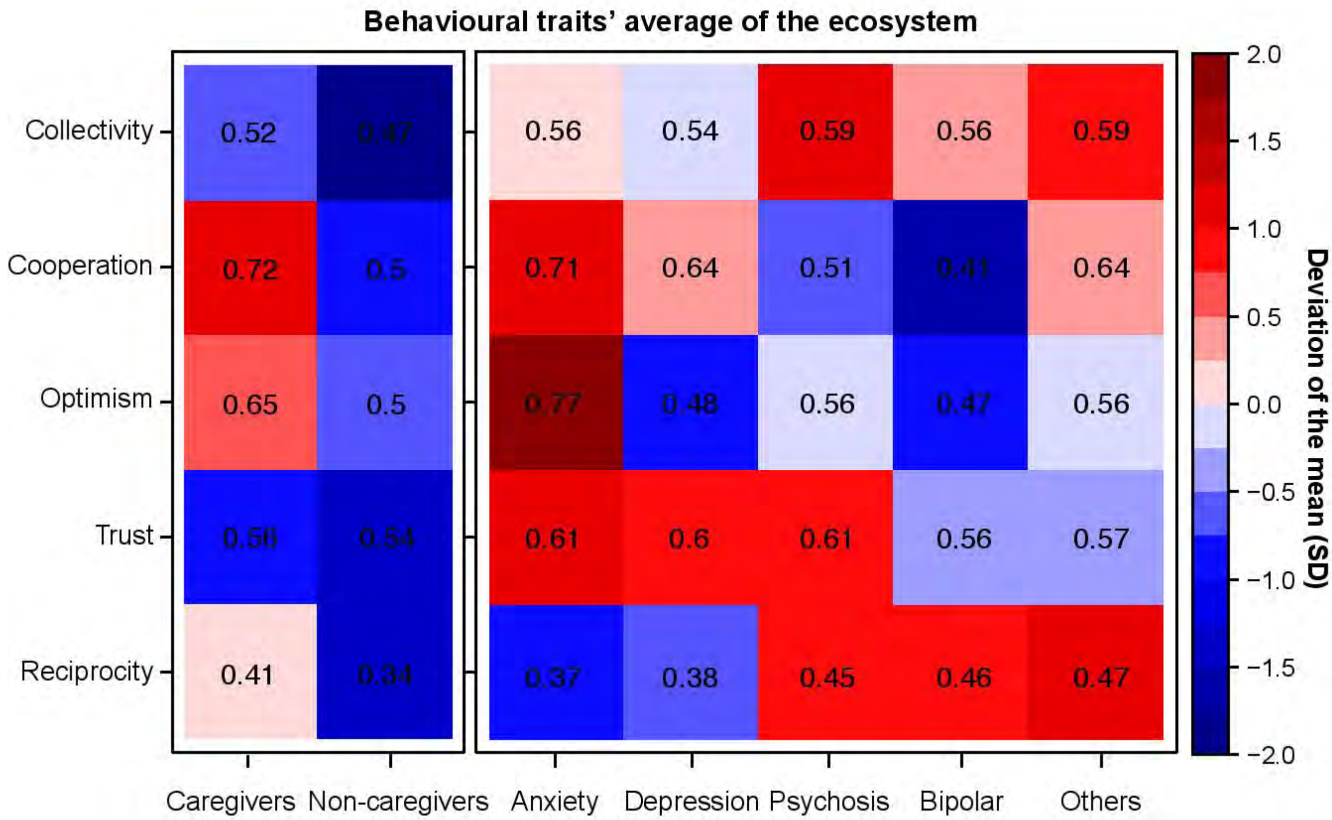


Figure 1. Heatmap of behavioural traits average and deviation of the mean across games. Collectivity refers to the ratio of contribution in the Collective-Risk Social Dilemma. Cooperation and Optimism refers to the ratio of cooperation and expected cooperation, respectively, in Prisoner's Dilemma. Trust and Reciprocity refers to the ratio of capital trusted and reciprocated in Trust Game. The left part shows the ratio of individuals without mental conditions: caregivers (professionals and relatives) and non-caregivers (relatives, friends and others). The right part shows the actions of individuals with mental conditions. Therefore, the number in each cell indicates the ratio of social preferences per subjects in each social dilemma and the color scale shows the deviation of the mean measured in SD units.

Participants with anxiety are also the ones with the most positive expectations about the partner's behavior compared to all but caregivers (see Supplementary Table 2). Also, relatives, friends and other members with no MD defect more than caregivers (Mann-Whitney U test, $U = 1352, p = 0.02839$), being relatives remarkably less cooperative than the rest of the collective $c = 0.33 \pm 0.16$. This suggests that cooperation among members of the mental health ecosystem is contextually based, depending on the role that actors play in the recovery process. It also varies across diagnostics, revealing a marked cooperativeness and optimism of individuals with anxiety disorders.

On the other hand, in sequential dyadic interactions (TG) all participants trust more than half of their endowment, being the distribution of initial transfers similar across groups. No variation is indeed found in trust levels in between participants with MD, caregivers and non caregivers (Kruskal-Wallis rank sum test, $H = 2.75, df = 2, p = 0.25$). Yet, at the time of reciprocating the partner's behavior, participants with anxiety and depression return the least ($37.5 \pm 3.3\%$). The difference is significant if compared to the return transfers of participants with psychosis or other diagnostics (see Supplementary Table 4).

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 results 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 the total amount needed for the threshold (120 MUs) equally 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 (See Supplementary Section 1.6.2).

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 (see Supplementary Section 1.6.1). Similarly to other public good experiments, contributions decrease over time[23]. 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 Figure 2).

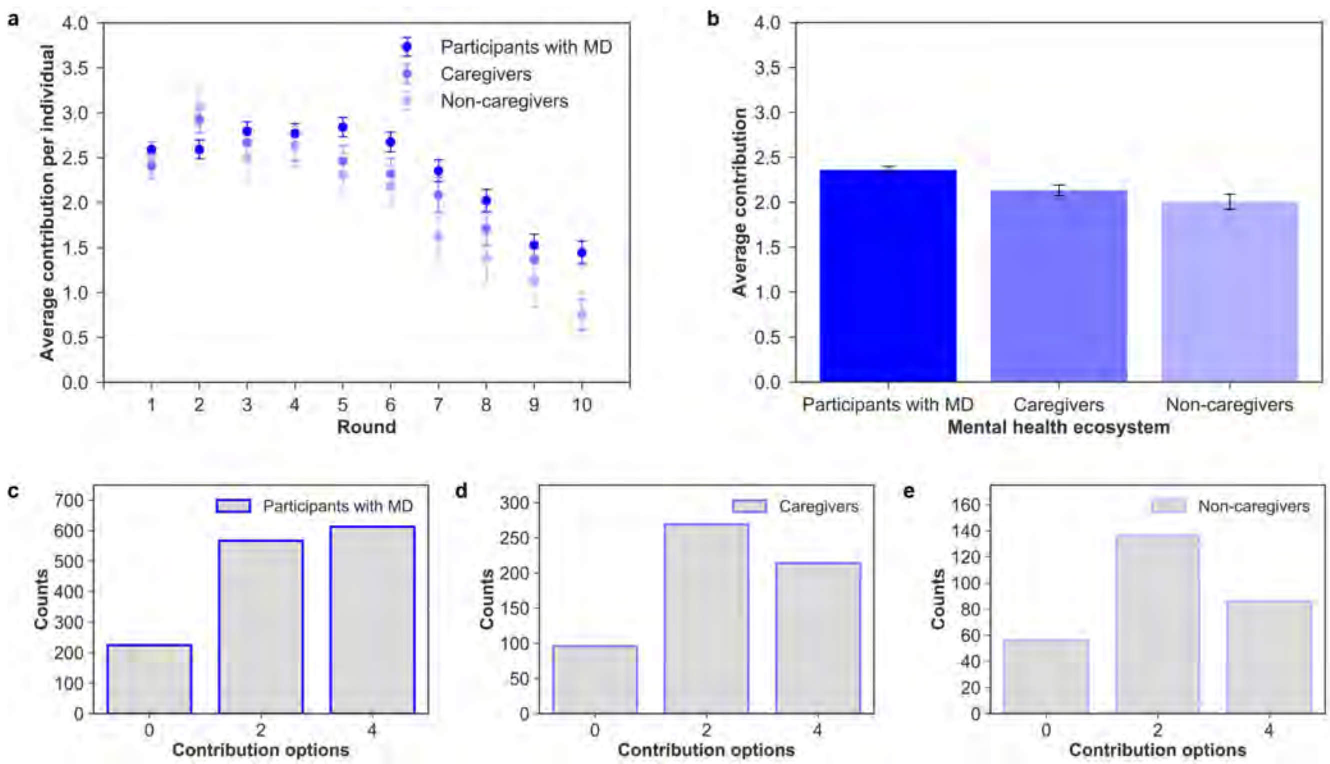


Figure 2. a: Individual contribution over rounds. Evolution of contributions (mean±sem) during the game between participants with mental disorder conditions, caregivers and non-caregivers. We can see that all groups behave similarly and in an identical way to a previous experiment run outside the mental health ecosystem[38]. **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 participants with MD (c), caregivers (d) and non-caregivers (e).** 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).

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 Figure 3b 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 achieved the goal, in average, later than in games with more than 50% of participants affected with MD (see Figure 3a).

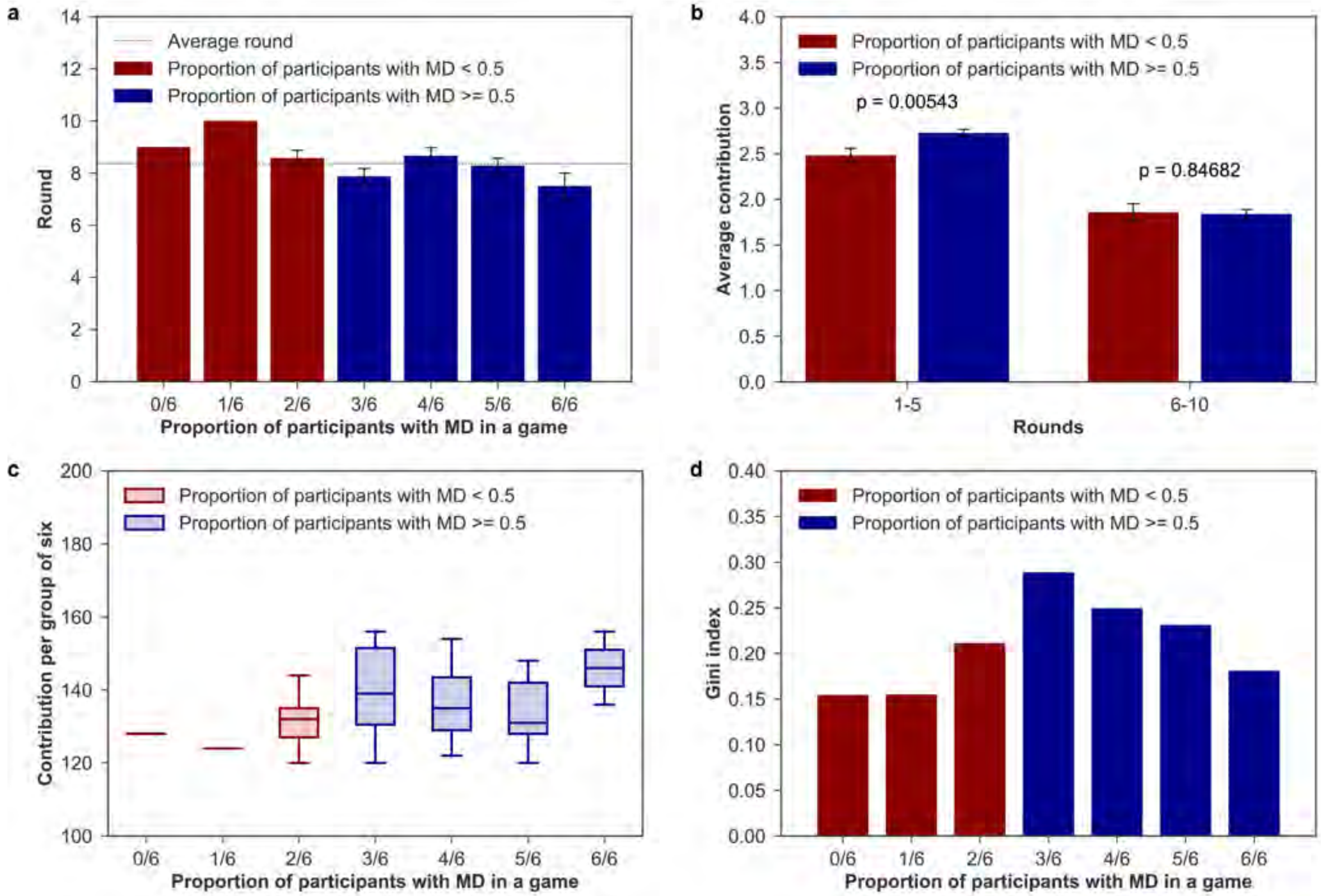


Figure 3. a: Average round of achievement. Round in which the group of six achieved the target. **b: Aggregated contributions per group composition.** Contributions in the first and last five-rounds per number of individuals with MD in a group. There are significant differences (t-test $p < 0.01$) in 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 mental disorder in each group.

If we then break down the analysis by group type, we find that group members contribute and benefit differently from cooperation (see Figure 3c). Indeed, final payoffs within groups are far from being equally distributed (see Figure 3d), with the highest inequality found in the group where the number of patients equals the number of actors involved in the recovery process (Gini 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, if in 4/6 and 5/6 groups caregivers reduce average individual contributions while non-caregivers pay more than their fair share. On the other hand, in 1/6 and 2/6 groups caregivers are the one 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.

Discussion

Let us now turn to the discussion of the above results and their implications (see Table 1 for a summary of the key findings). 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, participants with anxiety symptoms display a markedly different behavior compared to other diagnostics: they are more likely to opt for the cooperative strategy compared to individuals with bipolar disorder or depression, and return significantly less than participants with psychosis or other disorders. Since the current study is the first to investigate social decision-making within a heterogeneous population of individuals diagnosed with MD, a comparison with previous research is only possible referring to studies focusing on specific clinical and quite homogeneous populations. Several experiments have demonstrated deficits in cooperative behavior among individuals with anxiety or depression when playing iterated versions of the PD[11,17,24,25], but results about altruism (Ultimatum game) and trust are inconsistent between studies[6,7,11,12,17,26]. Individuals with major depressive disorders (which include anxiety and depressive symptoms) have also been found to systematically differ when their emotional responses to fairness are compared[6,17], showing higher levels of negative feelings when faced with unfair treatments. One of the hypothesis advanced to explain the systematic behavioral differences of individuals with anxiety relates to a potentiated sensitivity to negative stimuli as well as a tendency to treat neutral or ambiguous stimuli as negative or as less positive[6,12,17,27]. This hypothesis might find support in our results as for the low returns in the Trust Game, despite displaying relatively high trust in the partner's behavior and very high expectations. Indeed, participants with depressive or anxiety symptoms in our experiment significantly over-punish trustee transfers, but the low returns are independent of the amount received. This seems to imply that participants with mood disorders respond negatively to their partner behavior, as if they interpret their partner's choice in a negative sense. Alternatively, fairness considerations may be playing a role: low returns of participants with mood disorders might therefore be due to different fairness perceptions[6,12,17], which result in a bias towards negative reactions rather than positive rewarding.

Deficits in economic game play have also been documented for individuals with bipolar disorder. Studies report low and decreasing trust levels over sequential interactions, skeptical beliefs about the partner's behavior and a tendency to break cooperative interactions[28,29]. Again, this is partly supported by our results. Negative expectations of participants with bipolar disorder indeed agree with a low frequency of cooperative choices, little amounts of money sent to trustees, and low contributions to collective action. In line with King-Casas et al. results[29], while individuals with depression trust in the cooperativeness of other people, those with bipolar personality disorders do not. Cognitive dysfunctions (insula response) might possibly reflect an atypical social norm in this group[29]. Consequently, defection by partners might not violate the social expectations of individuals with BPD. In contrast, in our experiment, participants with bipolar disorder return the most within the group of individuals with a mental disorder. That is, they report a strong willingness to positively respond to a norm of trust as to signal their partner trustworthiness. Therefore, conditioned on the previous action of the partner, it seems that individuals with BPD are willing to show cooperative behavior. Considering now individuals with high levels of psychopathy, they have been found to make less fair offers, accept less fair offers, and show very high levels of defection[15,16,30]. Major explanations for such behavior point to deficits in emotion regulations (amygdala dysfunctions), which would lead to lack of anxiety, empathy, and guilt, coupled with exaggerated levels of anger and frustration[30] and to the absence of prepotent biases toward minimizing the distress of others[16]. In this case, our experiments do not confirm those previous results: Indeed, participants with psychosis are the ones who trust, contribute the most to the public good, and are willing to take costly actions to reciprocate their partner's behavior. It could be possible that, as psychopathic disorders are in fact a large group of different ones, behavioral differences among subgroups may lead to this discrepancy. In connection with these results, it is interesting to note that recent results on a large population of patients with paranoia suggest that distrust is not the best explanation for reduced cooperation and alternative explanations incorporating self-interest might be more relevant[31,32]. This calls for further research into this particular family of MD to clarify whether or not the behavioral characterization applies to all or to a subclass of them.

However, pointing to deficits in social cognition can only account for a partial explanation of individual behavior, and does not contribute to community care narratives. The fact that nothing in this direction has been reported before also reinforces the need to adopt a more holistic view on the interdependencies at play within the mental health collective. Indeed, if statistically relevant differences in cooperative behavior are found across diagnostics, they also depend on the role that actors play in the recovery process. That is, caregivers display exceedingly large degrees of cooperativeness and optimism in one-shot interactions. Caregivers can be thus considered the strong ties of the mental health ecosystem, of particular value when one seeks emotional support. With the de-institutionalization of health systems, caregivers have indeed become key players in care provision. Taking into account their behavior and expectations is therefore of particular interest to extend the support tailored to their needs. These actions should improve the effectiveness of their role by guiding them[33]. Yet, relatives who do not strictly contribute to caregiving practices turn out to be the weak links. It is thus likely that interventions designed to increase their participation in the community might help improve the recovery process.

Also, members of the mental health ecosystem do not equally contribute and 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. 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. Overall, the results point to the availability of large social capital in the mental health community that can make a difference in the welfare and recovery process of firsthand affected, and suggest that the community-centered approach to mental care may turn out to be very beneficial. 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. Though we depicted some behavioral differences in dyadic interactions, most importantly we found that individuals with MD show a remarkably larger disposition towards sustaining cooperation within groups. The results thus reinforce the idea of community social capital as a key approach to the recovery process based on an ecosystem paradigm (see also the recent results in Ref.[34] about the role and impact of family and community social capital on MD in children and adolescents). Also, if on the one hand the fact that the results of our dyadic games are in general agreement with previous studies validates our procedure; on the other hand it supports the validity and contributions of neuroeconomic and experimental approaches to the study of MD. 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. The adoption of our setup could lead to the identification of core groups that can boost and sustain cooperation within a given community. It can also help in discriminating among different communities in order to identify best practices and optimize resource allocation[35].

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.

Experimental design

As indicated in the main text, 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 experimental dilemmas being proposed served both to advance in knowledge on the social dynamics at play within 'care in the community' settings and as a self-reflection experience for all participants. The experimental design process developed in four main phases: (i) identification of the behavioral traits perceived as of fundamental importance within the community, (ii) operationalization of those same behavioral traits thorugh game theoretical paradigms and literature reviews, (iii) definition of the socio-demographic information relevant for the analysis, and (iv) a beta testing of the digital interface (including contents, time duration, and language used). 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 intermediaire between the scientists and different mental health collectives. It also provided valuable insights to better interpret the data obtained.

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 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 (see Supplementary Section 1.1). 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 (the experimental settings and instructions, can be found in the Supplementary Section 1.2 and 1.3 respectively). First, participants participated in a Collective Risk Dilemma[23] against five opponents. Briefly, the game is a public goods game with threshold: If the participants' total contribution after 10 rounds is lower than a given threshold, 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. After completing the task, participants played one round of the Trust Game[36] in both roles: as trustors and as trustees. They played against different partners in each role. Finally, they played one round of a Prisoner Dilemma[37] with (unincentivized) belief elicitation about their counterpart's behavior prior to playing. Before starting the games, participants had to complete a brief survey covering some key dimensions of their sociodemographic background. The assignment of players' partners in the dyadic games was completely random and every action was made with a different partner. The average (sem) time for completing the experiment (CRD, PD and TG) is around 4 minutes, $250.14 \pm 4.59s$. At the end of each session, participants received a gift card worth their earnings. The average individual earning is 46.84 ± 0.77 MUs equivalent to a 4.04 ± 0.077 EUR voucher. The behavioral patterns that emerged do not reveal significant variation across the different experiments, which may suggest that our results are robust to generalizations (see Supplementary Section 1.7).

Statistical analysis

Results were analyzed at two levels: first, we tested for behavioral differences between the whole group of individuals with a mental disorder compared to members of the mental health ecosystem; we then checked for systematic behavioral variation across diagnostics and role played in the recovery process. In one shot, two-person dyadic interactions we performed Mann-Whitney-U tests for independent groups to compare the distributions of cooperative choices (PD), and initial and back transfers (TG), between individuals with and without a mental condition. We then 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 mental disorder, 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.

Participants with anxiety display markedly different behavior compared to other diagnostics.	
• More likely to opt for the cooperative strategy compared to participants with bipolar disorder.	MWU test $p < 0.05$.
• Associated with the most positive expectations about the partner's behavior.	$c_{exp} = 0.77 \pm 0.08$. MWU tests comparisons across diagnostics $p < 0.05$.
• Show a significantly high frequency of cooperative interactions compared to individuals with bipolar disorder or depression.	$c = 0.71 \pm 0.08$ MWU tests comparisons $p < 0.05$ and $p < 0.1$.
• Return significantly less than participants with psychosis or other disorders.	$r = 0.37 \pm 0.05$. MWU test comparisons $p < 0.1$.
Cooperation depends on the role that actors play in the recovery process.	
• Significant differences in the frequency of cooperative interactions across role groups.	KW-RS test, $p < 0.05$.
• Caregivers contribute with large degrees of cooperativeness and optimism.	$c = 0.72 \pm 0.05$, $c_{exp} = 0.65 \pm 0.06$.
• Relatives are the weak links of the ecosystem.	$c = 0.33 \pm 0.16$, $c_{exp} = 0.44 \pm 0.18$.
Individuals with MD support the cost of collective action.	
• MD contribute more than caregivers and non caregivers to the public good.	Independent t-test $p < 0.005$. Average contributions: 22.95 ± 0.63 MUs, and 20.34 ± 0.68 MUs respectively.
• Groups with half or more MD do better in sustaining cooperation in the first rounds.	Independent t-test $p < 0.01$.
• Inequality in the distribution of final payoffs is at his maximum in groups where MD constitute half of the group.	Gini coefficient 0.289.

Table 1. Summary of the key findings from the suite of dyadic and repeated interactions among members of the mental health ecosystem. MWU: Mann-Whitney-U, KW-RS: Kruskal-Wallis rank sum, c: cooperation level, c_{exp}: expected cooperation level, r: return ratio in Trust Game, and MUs: monetary units.

References

1. Steel, Z. et al. The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013. Int. J. Epidemiol. **43**, 476–493 (2014).

2. Insel, T. R., Collins, P. Y. & Hyman, S. E. Darkness invisible: The hidden global costs of mental illness. Foreign Affairs **94**, (2015).

3. White, R. G. [Ed], Jain, S. [Ed], Orr, D. M. R. [Ed] & Read, U. M. [Ed]. The Palgrave handbook of sociocultural perspectives on global mental health. The Palgrave handbook of sociocultural perspectives on global mental health. (2017).

4. World Health Organisation. The World Health Report 2001: mental health, new understanding, new hope. World Health Report 1–169 (2001).

5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Diagnostic and Statistical Manual of Mental Disorders 4th edition TR. (2000).

6. Gradin, V. B. et al. Abnormal brain responses to social fairness in depression: an fMRI study using the Ultimatum Game. Psychol. Med. **45**, 1241–51 (2015).

7. Shao, R., Zhang, H. & Lee, T. M. C. The neural basis of social risky decision making in females with major depressive disorder. Neuropsychologia **67**, 100–10 (2015).

8. Guroglu, B., van den Bos, W., Rombouts, S. A. R. B. & Crone, E. A. Unfair? It depends: Neural correlates of fairness in social context. Soc. Cogn. Affect. Neurosci. **5**, 414–423 (2010).

9. Wang, Y., Yang, L.–Q., Li, S. & Zhou, Y. Game Theory Paradigm: A New Tool for Investigating Social Dysfunction in Major Depressive Disorders. Front. Psychiatry **6**, 128 (2015).

10. King–Casas, B. & Chiu, P. H. Understanding interpersonal function in psychiatric illness through multiplayer economic games. Biol. Psychiatry **72**, 119–125 (2012).

11. Pulcu, E. et al. Social–economical decision making in current and remitted major depression. Psychol. Med. **45**, 1–13 (2015).

12. Wang, Y. et al. Impaired social decision making in patients with major depressive disorder. BMC Psychiatry **14**, 18 (2014).

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

14. Agay, N., Kron, S., Carmel, Z., Mendlovic, S. & Levkovitz, Y. Ultimatum bargaining behavior of people affected by schizophrenia. Psychiatry Res. **157**, 39–46 (2008).

15. Mokros, A. et al. Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. J. Abnorm. Psychol. **117**, 406–413 (2008).

16. Rilling, J. K. et al. Neural Correlates of Social Cooperation and Non-Cooperation as a Function of Psychopathy. Biol. Psychiatry **61**, 1260–1271 (2007).

17. Scheele, D., Mihov, Y., Schwederski, O., Maier, W. & Hurlemann, R. A negative emotional and economic judgment bias in major depression. Eur. Arch. Psychiatry Clin. Neurosci. **263**, 675–683 (2013).

18. Putnam, R. D. Bowling Aalone: The Collapse and Revival of American Community. J. Democr. **6**, 1–18 (2000).

19. Ostrom, E., in Handbook of Social Capital: The troika of sociology, political science and economics, pp. 17–35 (2009).

20. Sagarra, O., Gutiérrez-Roig, M., Bonhoure, I. & Perelló, J. Citizen science practices for computational social science research: The conceptualization of pop-up experiments. Front. Phys. **3**, 93 (2016).

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

22. Brañas-Garza, P., Rodríguez-Lara, I., & Sánchez, A. Humans expect generosity. *Sci. Rep.* **7**, 42446.

23. Milinski, M., Sommerfeld, R.D., Krambeck, H.-J., Reed, F.A., & Marotzke, J. The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc. Nat'l. Acad. Sci. USA* **105**, 2291–2294 (2008).

24. Surbey, M. K. Adaptive significance of low levels of self-deception and cooperation in depression. Evol. Hum. Behav. **32**, 29–40 (2011).

25. Clark, C. B., Thorne, C. B., Hardy, S. & Cropsey, K. L. Cooperation and depressive symptoms. J. Affect. Disord. **150**, 1184–1187 (2013).

26. Destoop, M., Schrijvers, D., De Grave, C., Sabbe, B. & De Bruijn, E. R. A. Better to give than to take? Interactive social decision–making in severe major depressive disorder. J. Affect. Disord. **137**, 98–105 (2012).

27. Harlé, K. M., Allen, J. J. B., Sanfey, A. G., Harlé, K. M. & Sanfey, A. G. The impact of depression on social economic decision making. J. Abnorm. Psychol. **119**, 440–446 (2010).

28. Unoka, Z., Seres, I., Aspán, N., Bódi, N. & Kéri, S. Trust game reveals restricted interpersonal transactions in patients with borderline personality disorder. J. Pers. Disord. **23**, 399–409 (2009).

29. King-Casas, B. et al. The rupture and repair of cooperation in borderline personality disorder. Science **321**, 806–810 (2008).

30. Koenigs, M., Kruepke, M. & Newman, J. P. Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients. Neuropsychologia **48**, 2198–2204 (2010).

31. Raihani, N. J. & Bell, V. Paranoia and the social representation of others: a large-scale game theory approach. Scientific Reports **7**, 4544 (2017). https://doi.org/10.1038/s41598-017-04805-3

32. Raihani, N. J. & Bell, V. Conflict and cooperation in paranoia: a large-scale behavioural experiment. Psychological Medicine 1-11. doi:10.1017/S0033291717003075

33. Collins, L. G. ·& Swartz, K. Caregiver care. American family physician **83**, 1309–1317 (2011).

34. McPherson, K. E. et al. The association between social capital and mental health and behavioural problems in children and adolescents: an integrative systematic review. BMC psychology **2**, 7 (2014). https://doi.org/10.1186/2050-7283-2-7

35. Almedom, A. M. Social capital and mental health: An interdisciplinary review of primary evidence. Social Science and Medicine **61**, 943–964 (2005).

36. Berg, J., Dickhaut, J., & McCabe, K. Trust, reciprocity, and social history. Games Econ. Behav. **10**, 122–142 (1995).

37. Rapoport, A. & Chammah, A.M. Prisoner's Dilemma (University of Michigan Press, 1965).

38. Vicens, J. et al. Resource heterogeneity leads to unjust effort distribution in climate change mitigation. Preprint at https://arxiv.org/abs/1709.02857 (2017).

Acknowledgements

General: We thank the community of patients, caregivers and families working within the Federació de Salut Mental Catalunya (Catalonia Mental Health Federation) for the enthusiasm and for their invaluable help in the design and realization of the experiments. We are also specially thankful to I Bonhoure for the necessary logistics to make the experiments possible, to F Español for contributing in the first steps in the experimental design, to M Poll for always giving us the institutional support from inside the Federation, to both E Ferrer and F Muñoz for building the bridge between us and the Mental Health ecosystem and to X Trabadó for encouraging us to run this research.

Funding: This work was partially supported by Federació de Salut Mental Catalunya; by MINEICO (Spain), Agencia Estatal de Investigación (AEI) and Fondo Europeo de Desarrollo Regional (FEDER) through grants FIS2013-47532-C3-1-P (JD), FIS2016-78904-C3-1-P (JD), FIS2013-47532-C3-2-P (JP), FIS2016-78904-C3-2-P (JP); by Generalitat de Catalunya (Spain) through Complexity Lab Barcelona (contract no. 2014 SGR 608, JP) and through Secretaria d'Universitats i Recerca (contract no. 2013 DI 49, JD, JV); and by the EU through FET Open Project IBSEN (contract no. 662725, AS) and FET-Proactive Project DOLFINS (contract no. 640772, AS).

Author contributions statement

J.D., A.S., and J.P. conceived the original idea for the experiment; J.V and J.D. prepared the software for the final experimental setup, A.C. and J.V. analyzed the data; and all authors carried out the experiments, discussed the analysis results, and wrote the paper.

Additional information

Accession codes Data is available in an structured way at Zenodo public repository with DOI 10.5281/zenodo.XXXX.
Competing financial interests The authors declare no conflict of interest.

Quantitative account of social interactions in a mental health care ecosystem: cooperation, trust and collective action

Supplementary Information

Anna Cigarini,[1,2] Julián Vicens,[1,2,3] Jordi Duch,[3,8] Angel Sánchez,[4,5,6,7] and Josep Perelló[1,2,⋆]

[1]Departament de Física de la Matèria Condensada, Universitat de Barcelona, 08028 Barcelona, Spain

[2]Universitat de Barcelona Institute of Complex Systems UBICS, 08028 Barcelona, Spain

[3]Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007, Tarragona, Spain

[4]Grupo Interdisciplinar de Sistemas Complejos (GISC), Unidad de Matemática, Modelización y Ciencia Computacional, Universidad Carlos III de Madrid, 28911 Leganés, Spain

[5]Unidad Mixta Interdisciplinar de Comportamiento y Complejidad Social (UMICCS) UC3M-UV-UZ, Universidad Carlos III de Madrid, 28911 Leganés, Spain

[6]Institute UC3M-BS of Financial Big Data, Universidad Carlos III de Madrid, 28903 Getafe, Spain

[7]Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, 50009 Zaragoza, Spain

[8]Northwestern Institute on Complex Systems (NICO), Northwestern University, 60208 Evanston, IL, USA

November 13, 2017

Contents

1 Supplementary Notes

1.1 Sociodemographics

Supplementary Figure 1 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.

1.2 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 signaling 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.

1.3 Experimental instructions

Participants were given the following instructions (translated from the original Catalan and Spanish online versions).

1.3.1 Initial instructions

Screen 1

Welcome to Games for Mental Health

We propose you to participate in a game which is both an experience and a scientific inquiry. The data collected are part of a participatory research which combines together the interests of scientists, physicians, families, individuals with a mental condition, and social workers among others.

We want to understand with you how do we behave in a number of everyday situations. There is no right or wrong answer. Ah! Do not forget: You will be playing with other people, not with a robot.

Yours and the other player's personal interests will be at odds. As it does in number of decisions you make everyday when you head down the street. No more, no less.

Are you ready? It will only take 10 minutes.

Screen 2

Pick a pseudonym

You will need it at the end of the game to claim a prize, keep it in mind! The data collected are anonymous, the pseudonym prevents the identification of the players. Avoid whatever references to your personal identity.

Screen 3

Legal note

Personal anonymized data provided through the registration process and collected for Games for Mental Health project, will be included in the repository 'Proyectos de investigación, desarrollo e innovación con datos de carácter personal de nivel básico', property of the University of Barcelona. The data will be processed in accordance with the objectives set up in the aforementioned project. You can at any time exercise the rights of access, rectification and withdrawal from the project. For this purpose you are required to forward a written notification, together with a copy of your DNI or any other identity document, to Secretaria General de la Universidad de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona; or via email at secretaria.general@ub.edu.

I have read the above terms and conditions.

1.3.2 Sociodemographic questionnaire

Screen 4

Enter the following data

Question 1: *Why are you being looked after?* Answer 1: *Depression* Answer 2: *Bipolar disorder* Answer 3: *Psychosis (schizophrenia, etc.)* Answer 4: *Anxiety* Answer 5: *Other mental disorders* Answer 6: *I do not have any mental disorder*

Screen 4b

Enter the following data (IF Screen 4: 'I do not have any mental disorder')

Question 2: *Are you?* Answer 1: *Non-caregiving relative* Answer 2: *Caregiving relative or informal caregiver* Answer 3: *Friend* Answer 4: *Professional (health sector, social sector, etc.)* Answer 5: *Others*

Screen 5

Enter the following data

Question 3: *Employment status* Answer 1: *Not working* Answer 2: *Working* Answer 3: *Working in CET (Centro Especial de Trabajo)*

Question 4: *Gender* Answer 1: *Man* Answer 2: *Woman*

Question 5: *Enter your age*

Question 6: *Enter your Postal Code*

Screen 6

Enter the following data

Question 7: *Civil status* Answer 1: *Single* Answer 2: *Married* Answer 3: *Civil union* Answer 4: *Divorced* Answer 5: *Widow* Answer 6: *Others*

Question 8: *Educational level* Answer 1: *None* Answer 2: *Primary education* Answer 3: *Secondary education* Answer 4: *Baccalaureate or equivalent* Answer 5: *Professional education* Answer 6: *University studies* Answer 7: *Not specified*

1.3.3 Tutorial

Screen 7

Tutorial: How to play

(Screenshots in Supplementary Figure 10)

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.*

1.3.4 Game One: The Climate Game

(Screenshots in Supplementary Figure 11)

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. 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.

1.3.5 Game Two: Investor Game

(Screenshots in Supplementary Figure 12)

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?

You are: The investor!
You invested a total of...
And the entrepreneur returned you...
You earn...
And the entrepreneur earns...
Thank you very much for participating! Ready for the next game?

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?

You are: The entrepreneur!
The investor invested a total of...
And decided to return...
The investor earns...
And you earn...
Thank you very much for participating! Ready for the next game?

1.3.6 Game Three: Prize Game

(Screenshots in Supplementary Figure 13)

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 she earns 10.
If you choose KEEP and the other chooses MULTIPLY: You earn 5 and she earns 15.
If you choose MULTIPLY and the other chooses KEEP: You earn 15 and she earns 5.
If you choose MULTIPLY and the other chooses MULTIPLY: You earn 0 and she earns 0.
Is everything clear? If not, raise your hand.

Before playing the game... A CONFESSION: What do you think the other player will do?
Now it's your turn... What do you want to do with your prize?

You chose... You earn... She chose... She earns...

1.3.7 Results

Thank you for participating in all games!
Here is a summary of what you have earned in each game:
Climate Game:...
Investor Game:...
Entrepreneur Game:...
Prize Game:...
Total money:...
You have earned a voucher Abacus worth: ...

You can register at JocXlaSalutMental.org to receive some interesting preliminary results. Do not forget to claim your prize with your pseudonym.

1.4 Cooperation and Optimism in the Prisoner's Dilemma

Participants played a one-shot Prisoner's Dilemma (PD) with one random partner. They had to simultaneously choose to keep a prize (namely cooperate) or to multiply it (namely defect). No matter what the other does, defection yields a higher payoff than cooperation. The dilemma is that if both defect, both do worse than if both had cooperated. We define cooperation as the number of cooperative decisions over the total number of decisions. Before making the choice, participants were asked whether they believed the partner would cooperate or defect. Intuitively, believing that the opponent is going to cooperate involves hopefulness and confidence about the future, especially in interactions with uncertain consequences. We thus defined optimism as the number of times participants expected cooperation over the total number of expectations elicited. Figure 2 illustrates the joint distribution of participant cooperation and their beliefs about the partner's cooperation.

We tested whether estimates about the partner's decision do influence behaviour against the hypothesis that individuals act independently of what they expect other people do. We performed multiple Chi-squared tests to check for associations between expectations of cooperation and cooperative behaviour across groups. As Figure 2 illustrates, the relation turns out to be significant for individuals with MD (χ^2 = 17.442, p<0.0001) and caregivers (χ^2 = 7.6604, p = 0.005645) (Top), and for participants with psychosis (χ^2 = 8.422, p= 0.003707)(Bottom). We also compared all predictions with the observed partner's decision to see whether participants guess right. Overall, 69% of predictions are correct.

However, no significant differences are found in the frequency of cooperative interactions if we differentiate participants only by having or not a mental condition ($c = 0.58 \pm 0.04$ and $c = 0.65 \pm 0.05$ respectively). The distribution of cooperative choices for the two groups is indeed statistically equivalent (χ^2 = 1.143, p = 0.28). This also applies when participants' beliefs about the partner's behaviour are elicited (χ^2 = 0.12705, p = 0.72). Expecting the partner to cooperate or to defect does not significantly varies between participants with and without a mental condition ($c_{\text{exp}} = 0.57 \pm 0.04$ and $c_{\text{exp}} = 0.6 \pm 0.05$, respectively). Also, if cooperation varies across diagnostics and role groups, expectations do not. A Kruskal-Wallis test could not reject the null hypothesis of equality of distributions (Kruskal-Wallis, H= 2.3112, p = 0.3149). For further details about cooperation and expectation rates across groups see Table 1 and 2 respectively. For average cooperation and expectation across groups see Table 5.

1.5 Trust and Reciprocity in the Trust Game

Participants' played one round of the Trust Game in both roles: as senders and as returners, against different partners. They were given 10 Monetary Units (MUs) and had to decide how much money to send to their anonymous opponent between 0; 2; 4; 6; 8; or 10 MUs. Then, they had to decide how much of the total received to send back: either 0; 25%; 50%; 75% or 100% of the amount received and tripled by the experimenter.

The initial transfer captures player's intention to trust the opponent in the hopes of receiving the same amount of money (or more) in return. Yet, trusting involves risks in that the opponent has no obligation to reciprocate and to retain all the money earned for himself. Therefore, trusting may result in a smaller gain for the sender. We define trust as the amount of MUs transferred by the investor, normalized between 0 and 1. On the other hand, the back transfers of the returner provides with a baseline measure of reciprocity, which has to do with repaying in kind or rather punishing the partner's behavior. However, the decision to reciprocate trust also cause a reduction in the returner's gain. It thus depend on the offset between maximizing personal gains relative to the appreciation of the trust that was given. We defined reciprocity as the amount of MUs returned as a portion of the amount available to return. Participants were aware that they would play both roles with different partners.

Overall, most participants trust 6 MUs (31.8%) or 4 MUs (28.5%). Only 4 out of 270 participants send nothing, and as many as 15% transfer all their endowment to their anonymous opponent, the majority of whom (69%) has MD. The modal initial transfer of individuals with MD is 6 MUs trusted, while that of actors involved in the recovery process is 4 MUs. As for back transfers, most participants (41%) reciprocate with 50% MUs, that is they split what they received in half. However, as many as 39% participants return 25% of the amount available to return. Also, 7% participants keep all for themselves (67% of them belong to the MD group). Yet, 4% participants return everything, and almost all of them (92%) has a mental disorder. This implicitly suggest that untrustworthy behavior is more prevalent among participants with MD, yet they are more likely to return everything compared to caregivers and non caregivers. Further, while the modal behavior of individual with MD is to reciprocate with 25% MUs, the modal return transfer of individuals without MD is 50% MUs.

Supplementary Figure 3 aggregates results for individuals with MD, caregivers and non caregivers (Top) and for each diagnostic (Bottom). As it can be seen from the plot, on average, all participants trust more than half of their endowment, and the distribution of initial transfers is similar across groups. No variation is indeed found in trust levels between participants with MD, caregivers and non caregivers (Kruskal-Wallis, H= 2.7457, p= 0.2534). Also, all participants return, on average, positive amounts. However, some minor variation is found in back transfers across diagnostics, as mentioned in the letter, see Table 3 and 4. For average initial and back transfers across groups see Table 6.

1.6 Collectivity in the Collective-Risk Dilemma

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).

1.6.1 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 Supplementary Figure 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.

1.6.2 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 (see Supplementary Figure **??**) 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 Supplementary Figure 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).

1.6.3 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 (Supplementary Figure 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 (Supplementary Figure 7). In average, the contribution based on role doesn't differ depending of the composition of the group.

1.6.4 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.

1.6.5 Earnings and Contributions

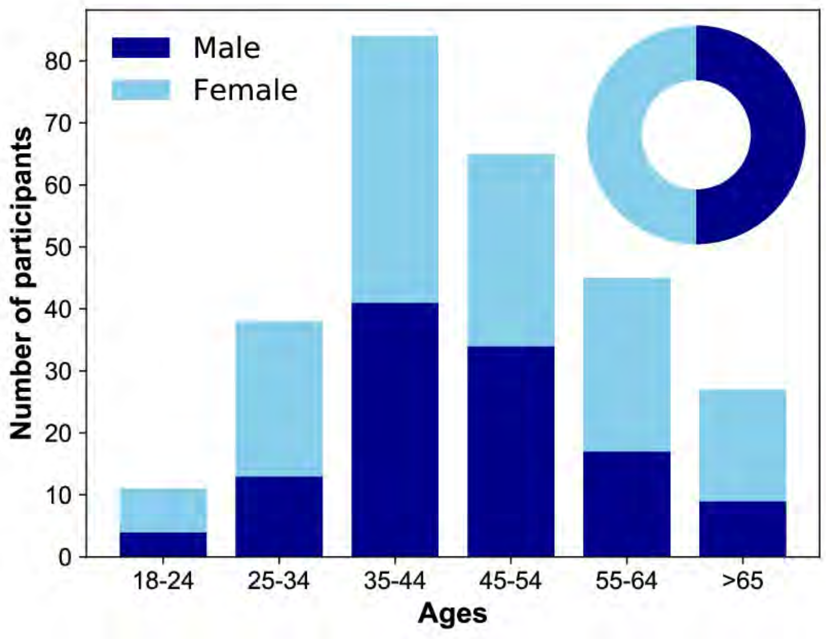
Final earnings include all participants (270). However, participants who did not contribute in two or more rounds, and had the computer contributing for them, did not get any profit. Their final earning is 0. The average (sem) earnings of all subjects is 16.50 ± 0.53 MUs (see Supplementary Figure 9 for final earnings across groups).

Individual contributions per round (Supplementary Figure 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.

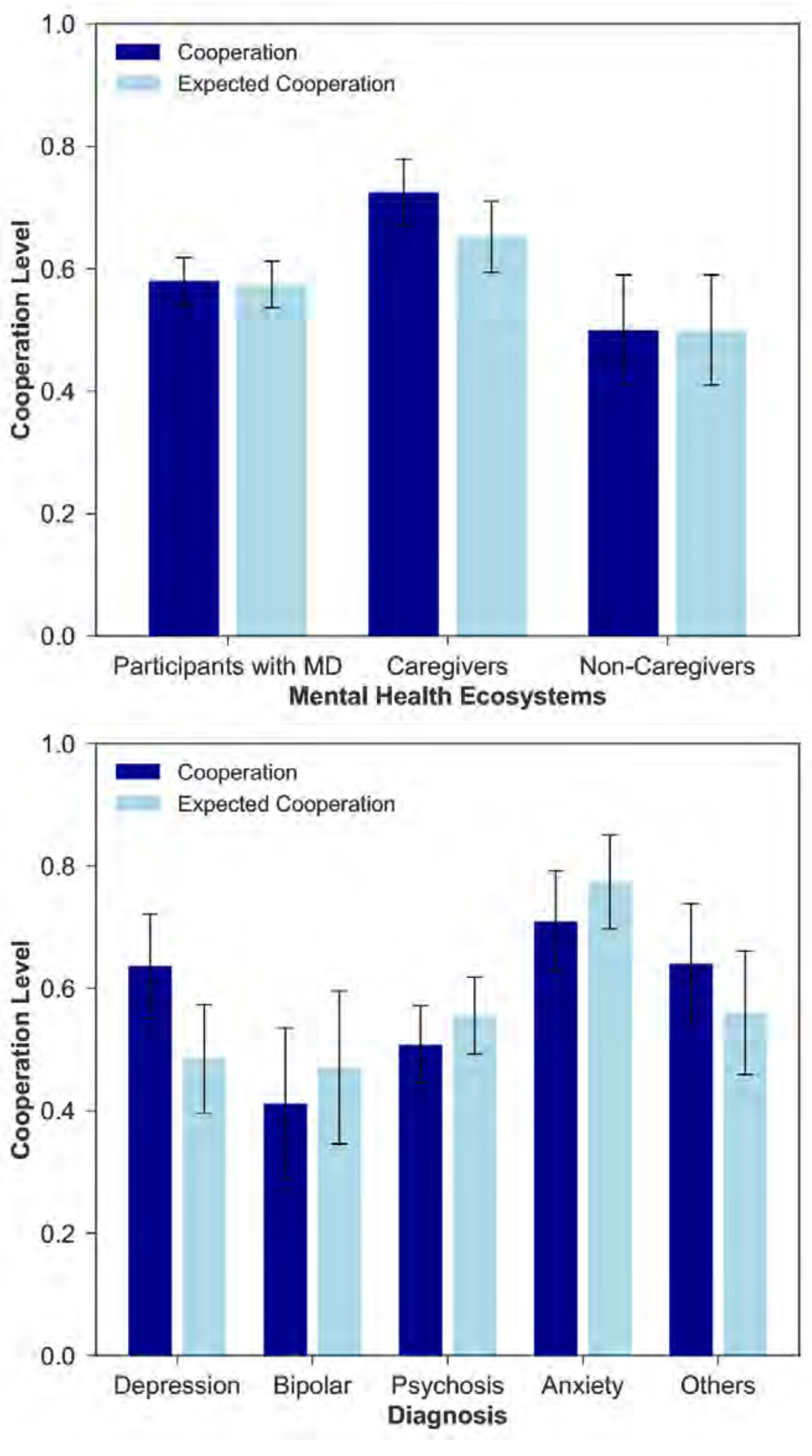
1.7 Robustness to generalizations

We tested for differences in game behaviors across the four experiments, namely Lleida, Girona, Sabadell and Valls experiments. No significant variation was found in the frequency of cooperative behaviors (Kruskal-Wallis, H= 2.3827, p = 0.4969), general expectations of cooperative behavior (Kruskal-Wallis, H= 0.37976, p = 0.9444), initial and back transfers in the TG (Kruskal-Wallis, H= 2.6683, p= 0.4456, and H= 3.0222, p= 0.3882), nor contributions to the common good (ANOVA, F= 0.21, p= 0.647).

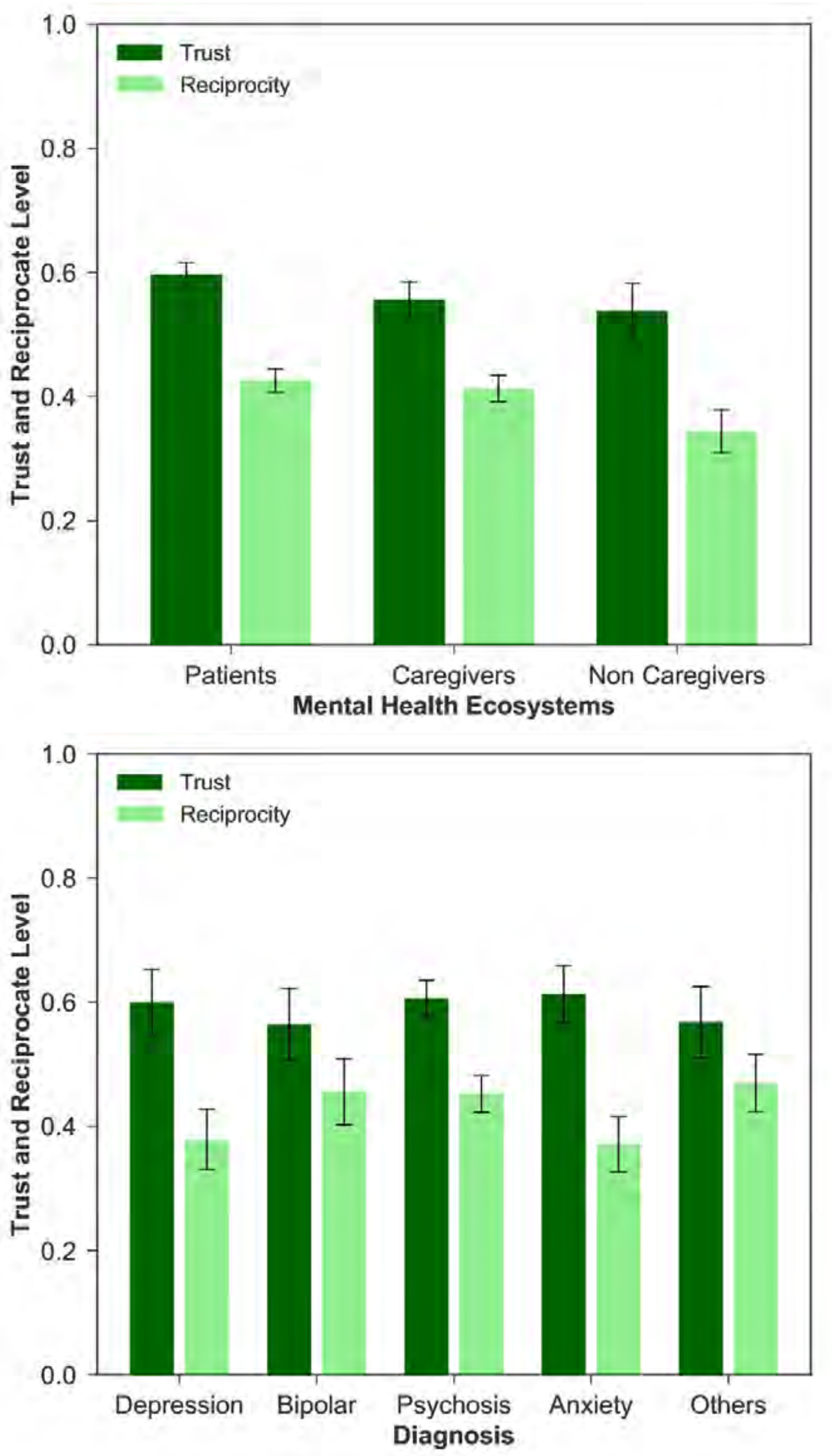
2 Supplementary Figures



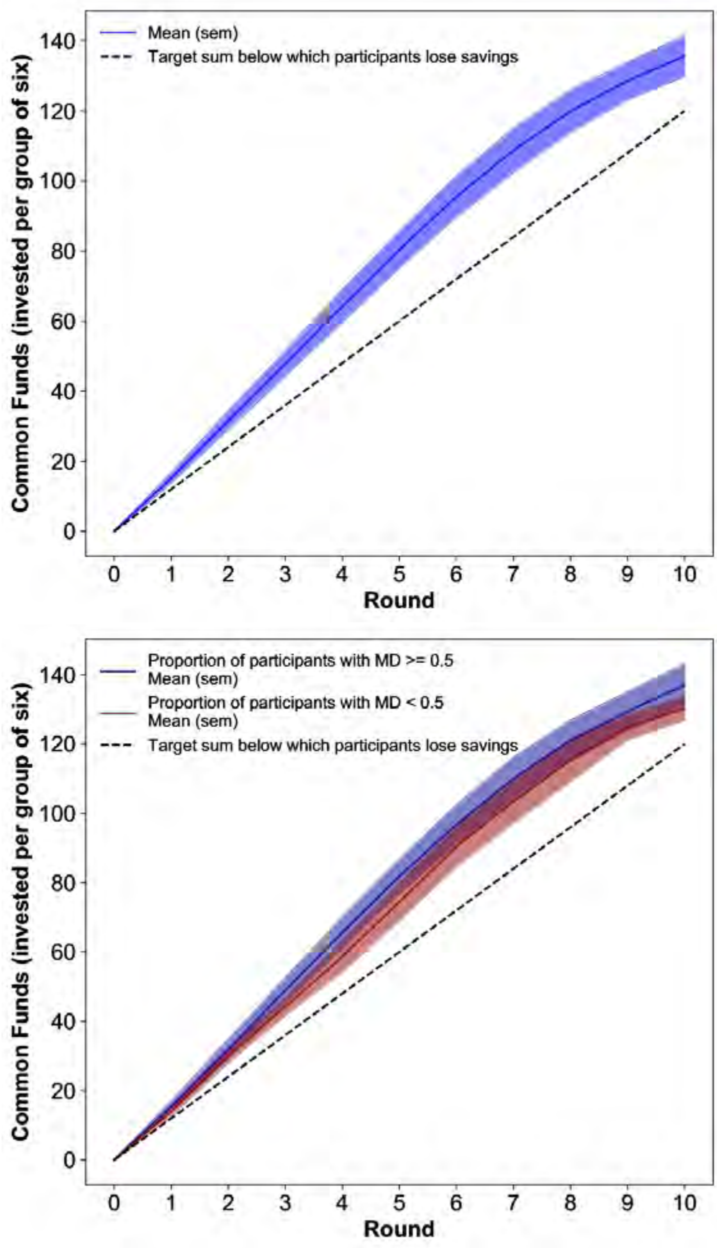
Supplementary Figure 1: Sociodemographics. Age and gender distributions. Age ranges are those given in the survey following Ethics and Privacy committe advice. There were 270 participants: 55.6% were men and 44.4% were women.



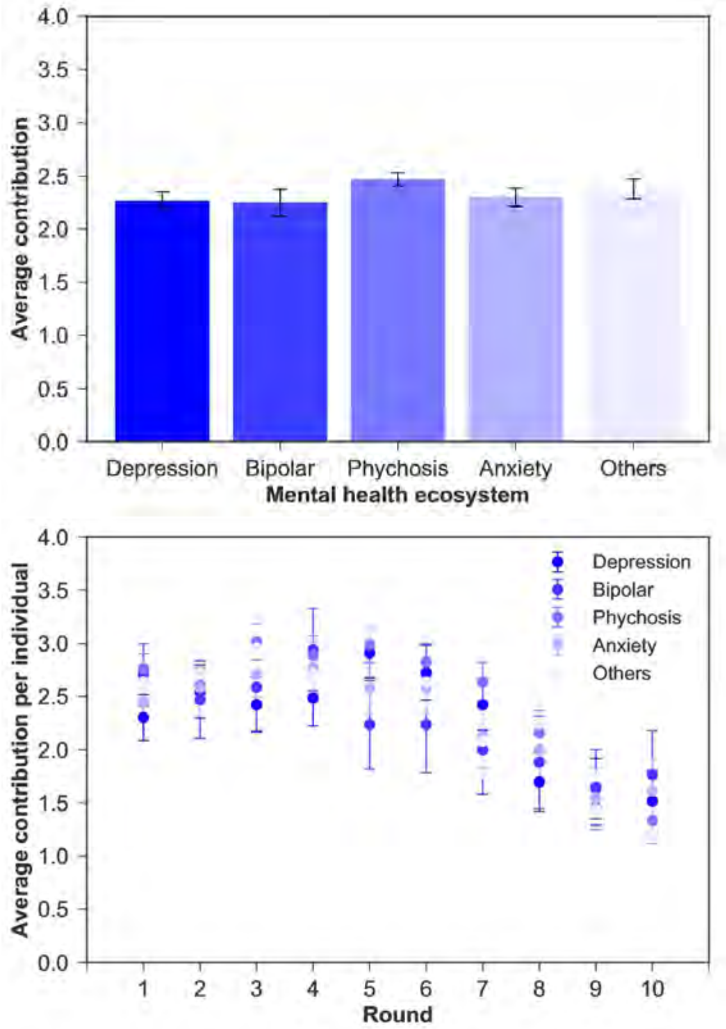
Supplementary Figure 2: Association between cooperation and expected cooperation in the Prisoner's Dilemma. The figure displays the frequency of cooperative choices and expected cooperation across groups (Top) and diagnostics (Bottom). Chi-squared tests for association between expectation and cooperation are significant for individuals with MD ($\chi^2 = 17.442$, $p<0.0001$) and caregivers ($\chi^2 = 7.6604$, $p < 0.01$) (Top), and for participants with psychosis ($\chi^2 = 8.422$, $p < 0.005$)(Bottom).



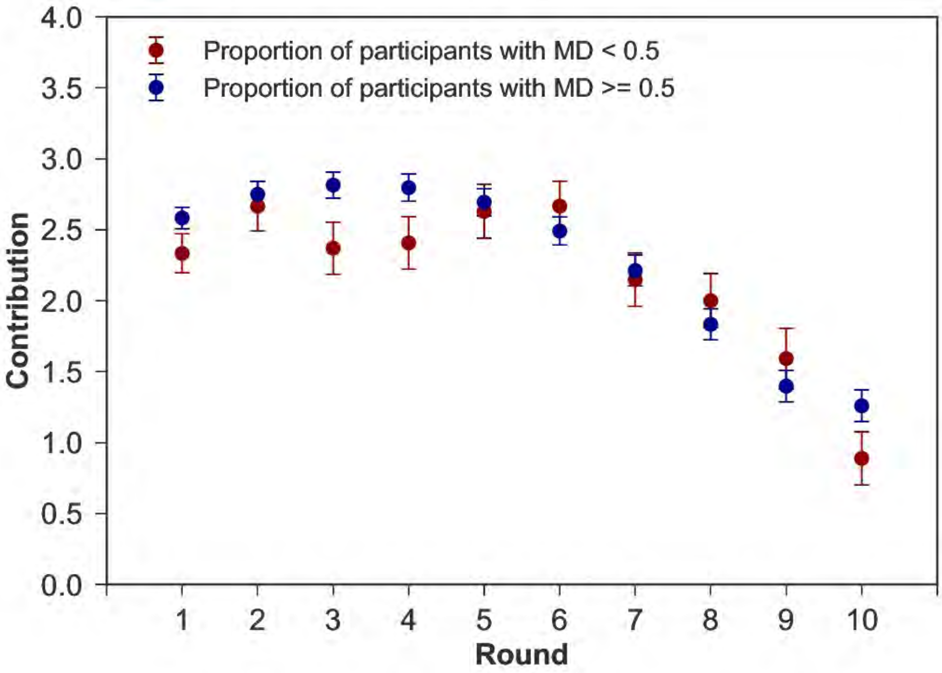
Supplementary Figure 3: Trust and Reciprocity levels in the Trust Game. Initial and back transfers in the TG across groups (Top) and diagnostics (Bottom).



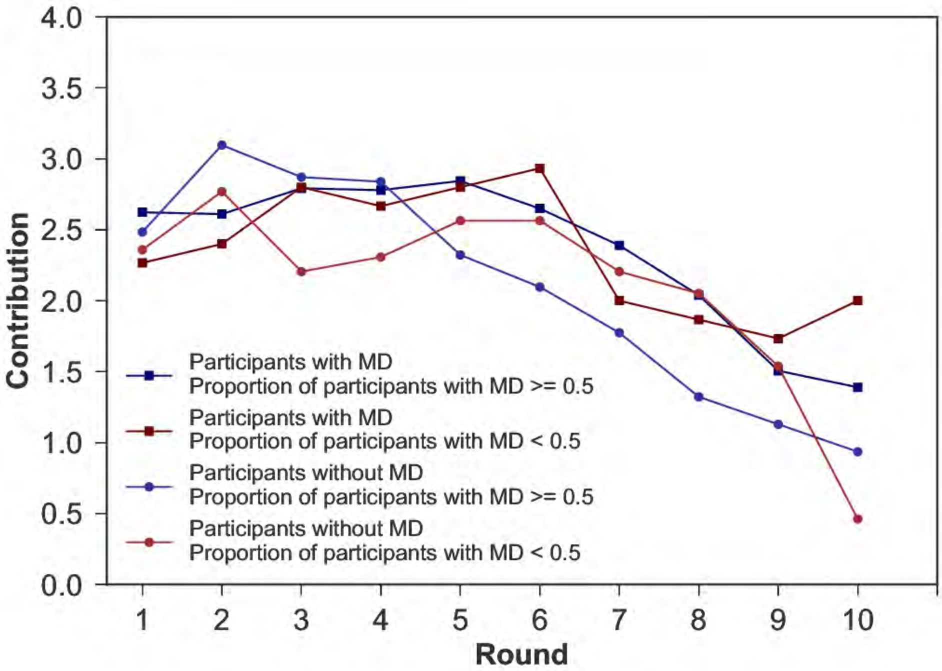
Supplementary Figure 4: Evolution of contributions during the games. (Top) Evolution of aggregate contributions to the common fund over rounds. (Bottom) Evolution of aggregate contributions over rounds depending on the portion of firsthand affected within groups.



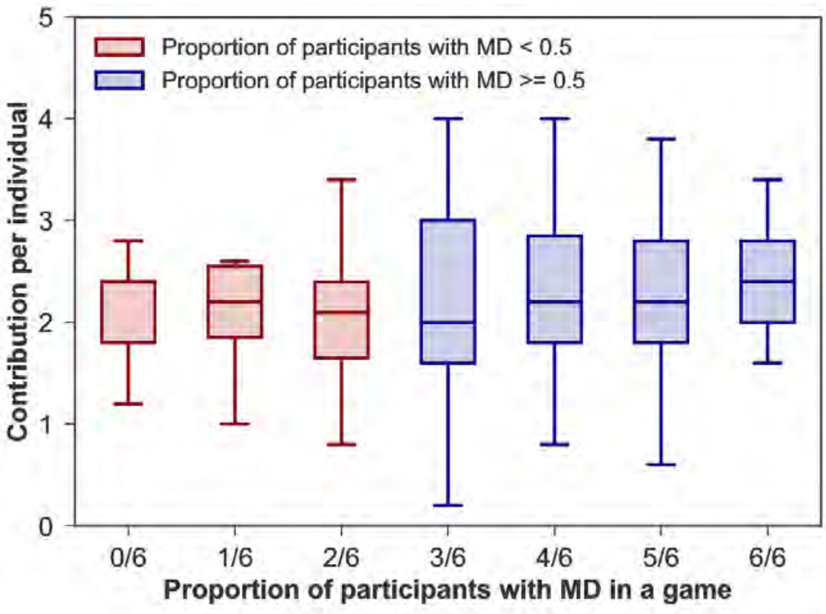
Supplementary Figure 5: Average aggregate contributions (Top) and evolution of contributions over rounds (Bottom) across diagnostics.



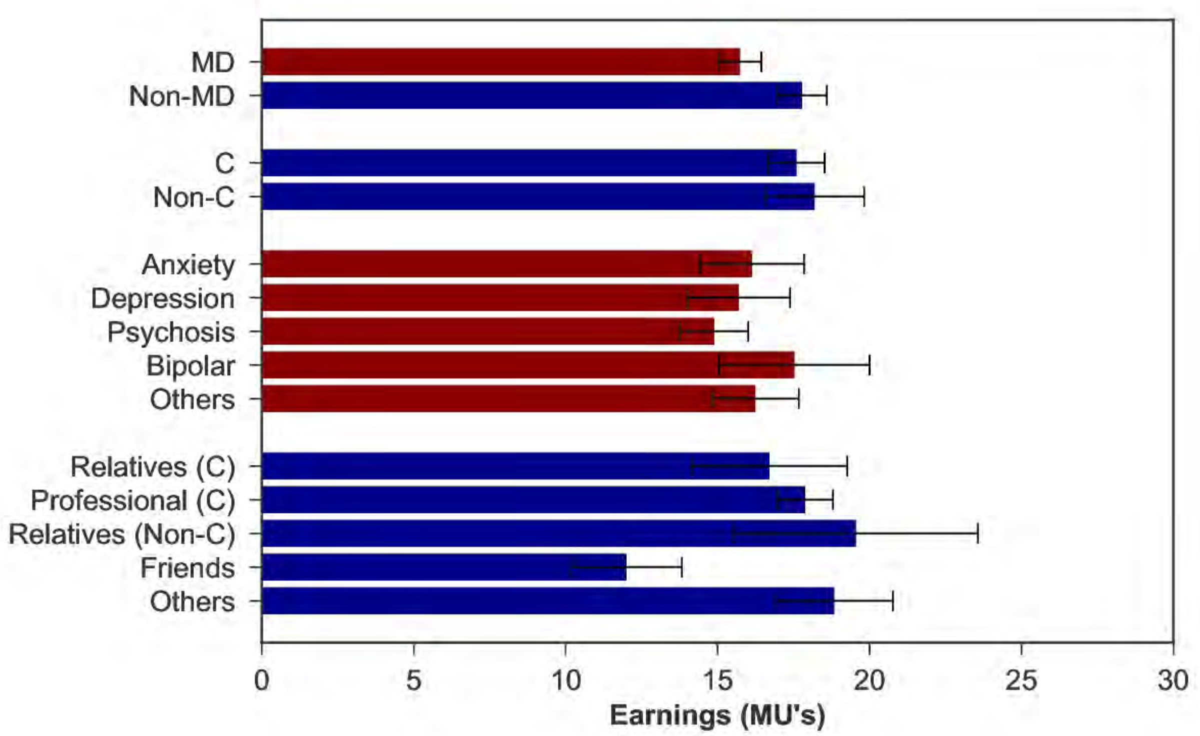
Supplementary Figure 6: Evolution of contributions over round according to the portion of patients within groups.



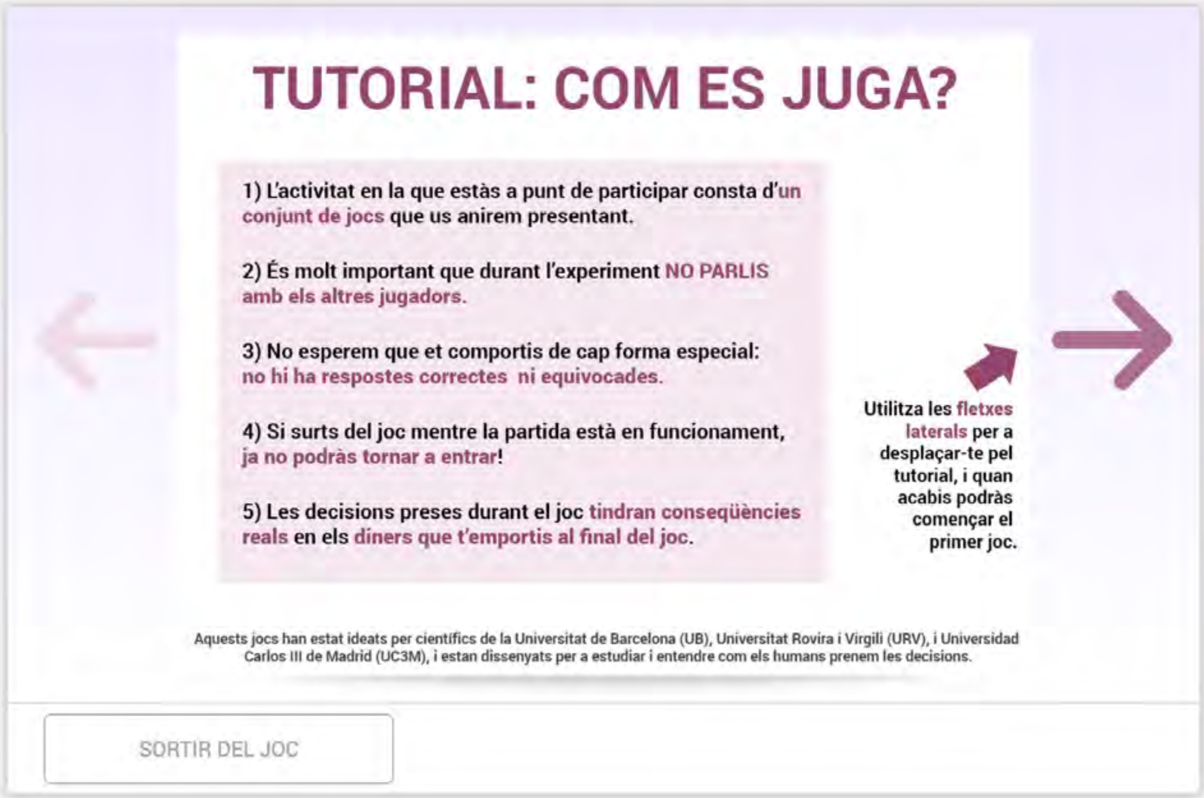
Supplementary Figure 7: Contributions over round by group composition. Evolution of individual contributions over round of individuals with and without a mental condition by portion of individuals with MD within groups.



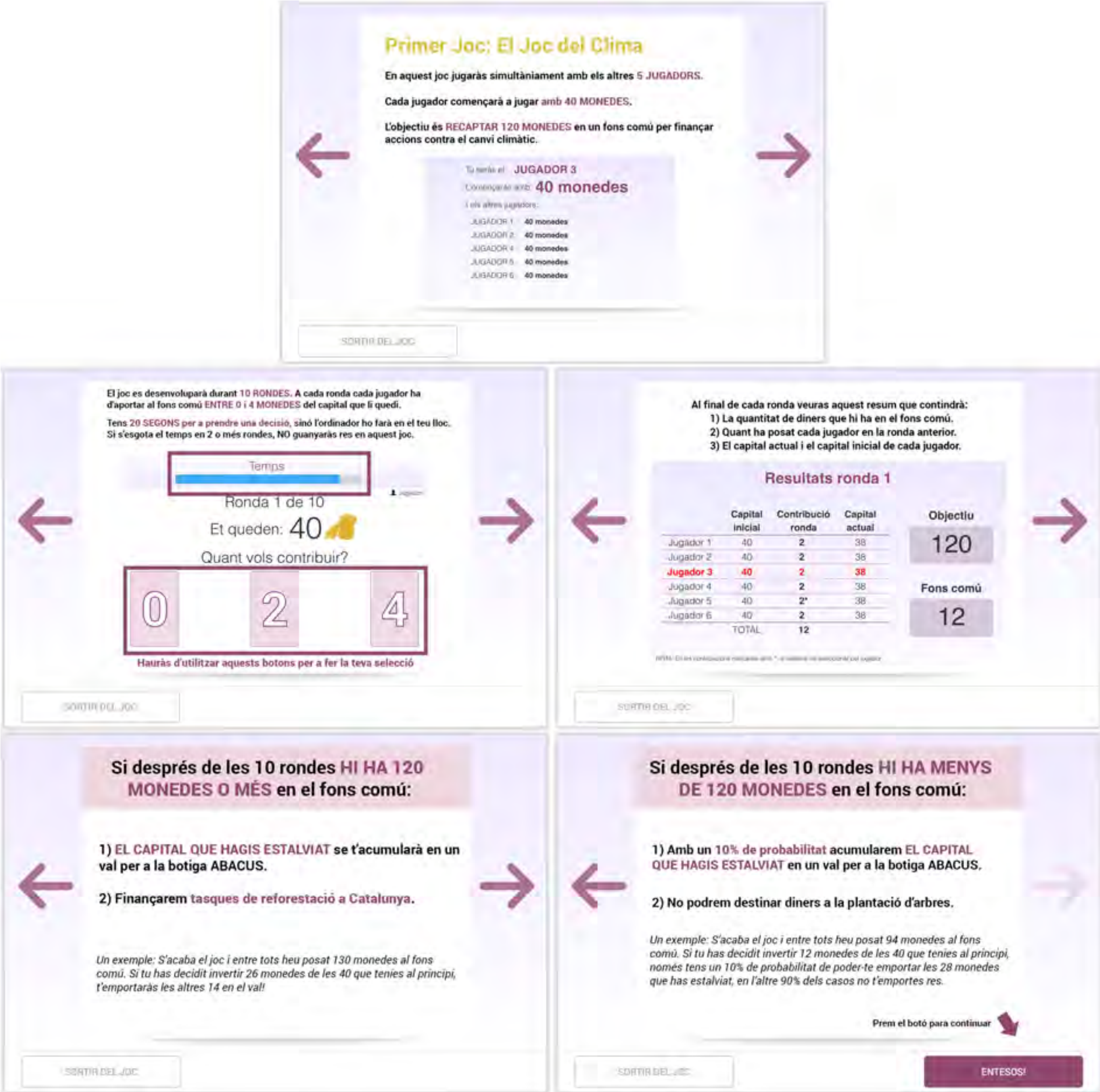
Supplementary Figure 8: Individual contribution. There are not significant differences in individual contribution by groups composition (ANOVA, F: 0.371 p: 0.9).



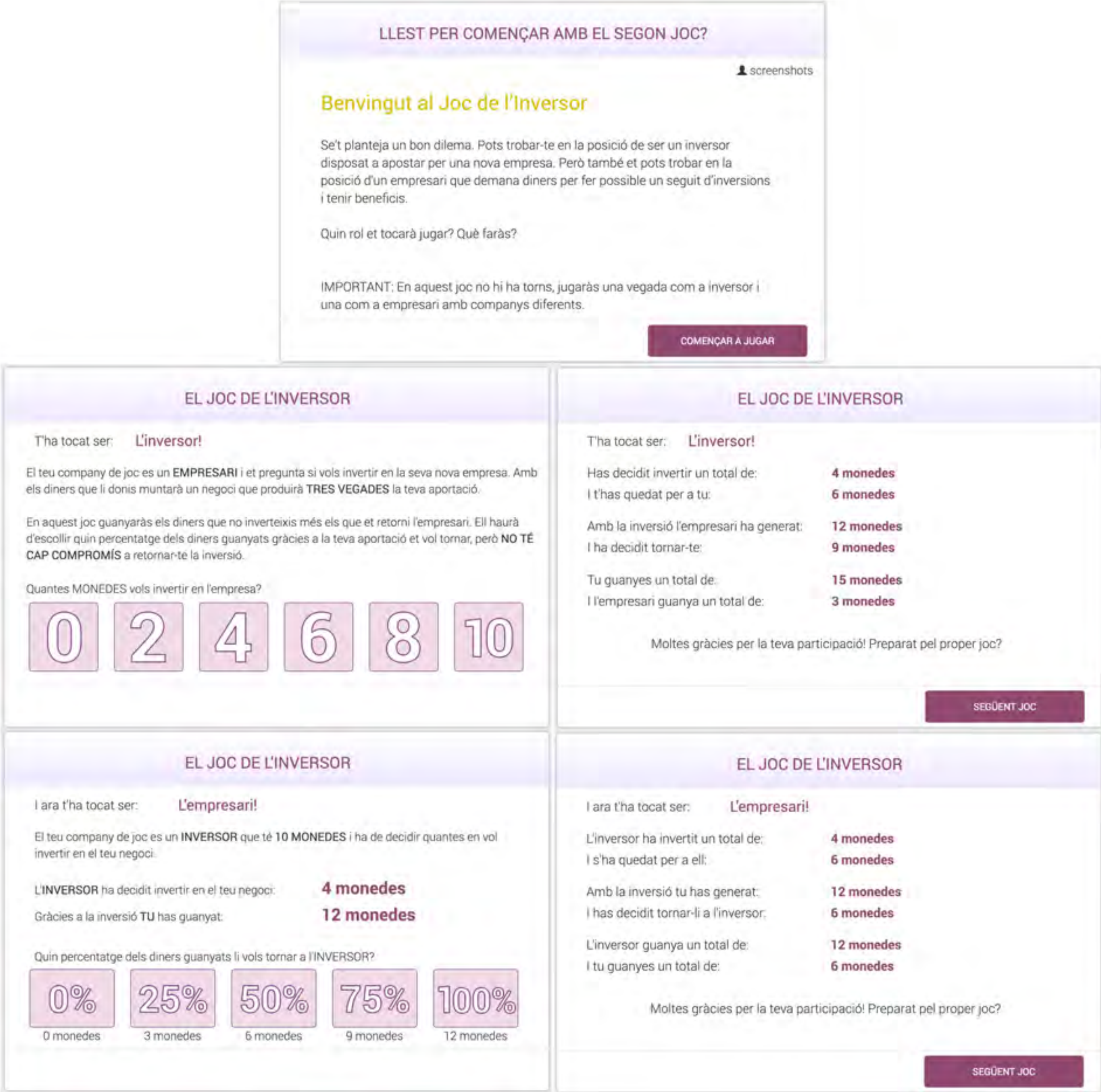
Supplementary Figure 9: Earnings in Colletive-Risk Social Dilemma. Earnings by role in the ecosystem. We show results from participants with and without Mental Disorder condition (MD and Non-MD, respectively), caregivers and non caregivers (C and Non-C, respectively), MD individuals with different diagnosis, and finally other actors that may and may not be caregivers.



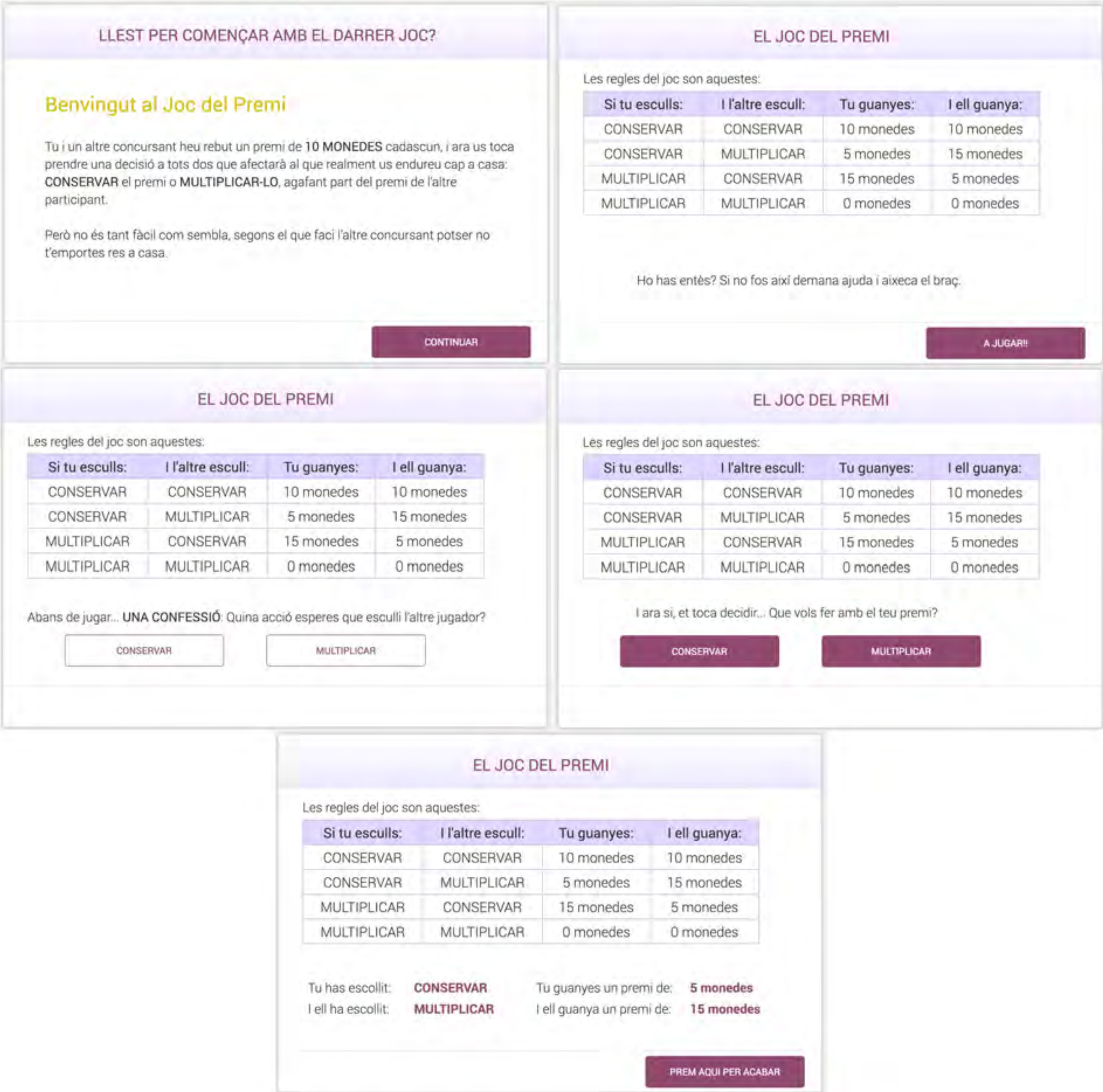
Supplementary Figure 10: Tutorial: How to play.



Supplementary Figure 11: Climate Game tutorial and game screenshots.



Supplementary Figure 12: Investor Game tutorial and game screenshots.



Supplementary Figure 13: Prize Game tutorial and game screenshots.

3 Supplementary Tables

Supplementary Table 1: Pairwise comparisons of cooperative behavior across diagnostics.

	Bipolar	**Pychosis**	**Depression**	**Anxiety**	**Others**
Bipolar		484 (0.48)	343.5 (0.13)	342 (0.045)**	261 (0.15)
Pychosis			1173 (0.23)	1173.5 (0.06)*	891.5 (0.26)
Depression				549 (0.53)	414 (0.97)
Anxiety					360.5 (0.58)
Others					

The table display the values of the Mann-Whitney-U tests (p-value) used to compare the frequency of cooperative decisions between each two diagnostics . *significant at 10%; **significant at 5%; ***significant at 1%.

Supplementary Table 2: Pairwise comparisons of expectations across diagnostics.

	Bipolar	**Pychosis**	**Depression**	**Anxiety**	**Others**
Bipolar		490 (0.54)	284.5 (0.92)	343.5 (0.03)**	231.5 (0.57)
Pychosis			966 (0.51)	1190 (0.04)**	791 (0.97)
Depression				659.5 (0.02)**	443.5 (0.57)
Anxiety					304.5 (0.09)*
Others					

The table display the values of the Mann-Whitney-U tests (p-value) used to compare the frequency of expected cooperation between each two diagnostics . *significant at 10%; **significant at 5%; ***significant at 1%.

Supplementary Table 3: Pairwise comparisons of trust across diagnostics.

	Bipolar	Pychosis	Depression	Anxiety	Others
Bipolar		466.5 (0.4)	311 (0.52)	297.5 (0.45)	217 (0.9)
Pychosis			1042.5 (0.98)	986.5 (0.93)	713.5 (0.48)
Depression				516.5 (0.94)	380 (0.6)
Anxiety					349.5 (0.52)
Others					

The table display the values of the Mann-Whitney-U tests (p-value) used to compare the initial transfers between each two diagnostics . *significant at 10%; **significant at 5%; ***significant at 1%.

Supplementary Table 4: Pairwise comparisons of reciprocity across diagnostics.

	Bipolar	Pychosis	Depression	Anxiety	Others
Bipolar		554 (0.81)	218 (0.18)	197.5 (0.13)	206 (0.86)
Pychosis			835.5 (0.097)*	761.5 (0.067)*	798 (0.92)
Depression				508.5 (0.97)	508 (0.11)
Anxiety					488 (0.07)*
Others					

The table display the values of the Mann-Whitney-U tests (p-value) used to compare the back transfers between each two diagnostics . *significant at 10%; **significant at 5%; ***significant at 1%.

Supplementary Table 5: Cooperation (c) and Expected Cooperation (c_{exp}) in Prisoner's Dilemma

<table>
<tr><td colspan="10">Experimental individuals
$n = 270$</td></tr>
<tr><td colspan="10">$c = 0.61 \pm 0.03$
$c_{exp} = 0.59 \pm 0.03$</td></tr>
<tr><td colspan="5">Non-MD
$n = 101$</td><td colspan="5">MD
$n = 169$</td></tr>
<tr><td colspan="5">$c = 0.65 \pm 0.05$
$c_{exp} = 0.60 \pm 0.05$</td><td colspan="5">$c = 0.58 \pm 0.04$
$c_{exp} = 0.57 \pm 0.04$</td></tr>
<tr><td colspan="2">Caregivers
$n = 69$</td><td colspan="3">Non-Caregivers
$n = 32$</td><td>Anxiety
$n = 31$</td><td>Depression
$n = 33$</td><td>Psychosis
$n = 63$</td><td>Bipolar
$n = 17$</td><td>Others
$n = 25$</td></tr>
<tr><td colspan="2">$c = 0.72 \pm 0.05$
$c_{exp} = 0.65 \pm 0.06$</td><td colspan="3">$c = 0.5 \pm 0.09$
$c_{exp} = 0.5 \pm 0.09$</td><td>$c = 0.71 \pm 0.08$
$c_{exp} = 0.77 \pm 0.08$</td><td>$c = 0.64 \pm 0.09$
$c_{exp} = 0.48 \pm 0.09$</td><td>$c = 0.51 \pm 0.06$
$c_{exp} = 0.55 \pm 0.06$</td><td>$c = 0.41 \pm 0.12$
$c_{exp} = 0.47 \pm 0.12$</td><td>$c = 0.64 \pm 0.09$
$c_{exp} = 0.56 \pm 0.1$</td></tr>
<tr><td>Professional
$n = 52$</td><td>Relatives
$n = 17$</td><td>Relatives
$n = 9$</td><td>Friends
$n = 4$</td><td>Others
$n = 19$</td><td colspan="5"></td></tr>
<tr><td>$c = 0.67 \pm 0.07$
$c_{exp} = 0.63 \pm 0.07$</td><td>$c = 0.88 \pm 0.08$
$c_{exp} = 0.71 \pm 0.11$</td><td>$c = 0.33 \pm 0.16$
$c_{exp} = 0.44 \pm 0.18$</td><td>$c = 0.25 \pm 0.25$
$c_{exp} = 0.25 \pm 0.25$</td><td>$c = 0.63 \pm 0.11$
$c_{exp} = 0.59 \pm 0.12$</td><td colspan="5"></td></tr>
</table>

Supplementary Table 6: Trust (t) and Reciprocity (r) in Trust Game

<table>
<tr><td colspan="10">Experimental individuals
$n = 270$</td></tr>
<tr><td colspan="10">$t = 0.58 \pm 0.01$
$r = 0.41 \pm 0.01$</td></tr>
<tr><td colspan="5">Non-MD
$n = 101$</td><td colspan="5">MD
$n = 169$</td></tr>
<tr><td colspan="5">$t = 0.55 \pm 0.02$
$r = 0.39 \pm 0.02$</td><td colspan="5">$t = 0.6 \pm 0.02$
$r = 0.43 \pm 0.02$</td></tr>
<tr><td colspan="2">Caregivers
$n = 69$</td><td colspan="3">Non-Caregivers
$n = 32$</td><td>Anxiety
$n = 31$</td><td>Depression
$n = 33$</td><td>Psychosis
$n = 63$</td><td>Bipolar
$n = 17$</td><td>Others
$n = 25$</td></tr>
<tr><td colspan="2">$t = 0.56 \pm 0.03$
$r = 0.41 \pm 0.02$</td><td colspan="3">$t = 0.54 \pm 0.04$
$r = 0.34 \pm 0.03$</td><td>$t = 0.61 \pm 0.05$
$r = 0.37 \pm 0.05$</td><td>$t = 0.6 \pm 0.05$
$r = 0.38 \pm 0.05$</td><td>$t = 0.61 \pm 0.03$
$r = 0.45 \pm 0.03$</td><td>$t = 0.56 \pm 0.06$
$r = 0.46 \pm 0.05$</td><td>$t = 0.57 \pm 0.06$
$r = 0.47 \pm 0.05$</td></tr>
<tr><td>Professional
$n = 52$</td><td>Relatives
$n = 17$</td><td>Relatives
$n = 9$</td><td>Friends
$n = 4$</td><td>Others
$n = 19$</td><td colspan="5"></td></tr>
<tr><td>$t = 0.57 \pm 0.03$
$r = 0.42 \pm 0.03$</td><td>$t = 0.51 \pm 0.05$
$r = 0.4 \pm 0.03$</td><td>$t = 0.46 \pm 0.06$
$r = 0.44 \pm 0.07$</td><td>$t = 0.7 \pm 0.06$
$r = 0.25 \pm 0.1$</td><td>$t = 0.54 \pm 0.07$
$r = 0.32 \pm 0.04$</td><td colspan="5"></td></tr>
</table>