A new agent-based model offers insight into population-wide adoption of prosocial common-pool behavior

SCHOLARONE™
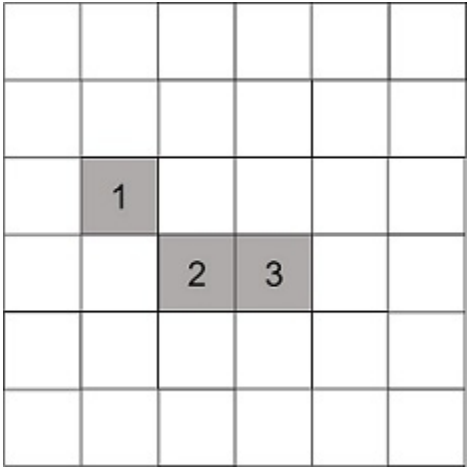Manuscripts

The Journal of Mathematical Sociology

Figure 1: An example-based explanation of how groups are defined in MGS I. The 6x6 landscape includes three agents that are organized into three groups. Agent 1's group consists of agent 2 and itself. Agent 2's group consists of agent 1, 3, and itself. Agent 3's group consists of agent 2 and itself.

39x39mm (150 x 150 DPI)

	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

Table 1: The payoff structure faced by individuals in a Prisoner's Dilemma, where T (temptation) > R (reward) > P (punishment) > S (sucker's).

110x44mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
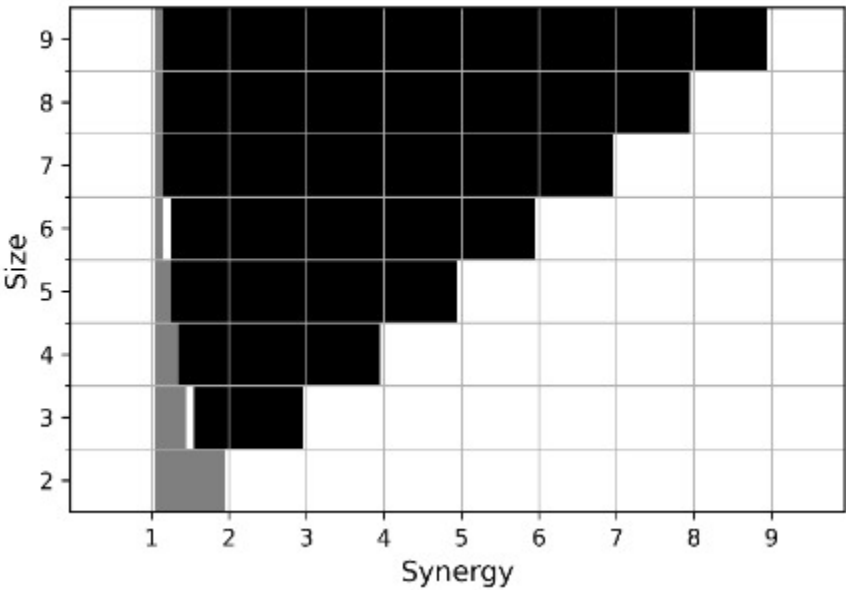11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Figure 2: The (group) size-synergy (from prosocial common-pool behavior) combinations in which agents can face a Prisoner's-Dilemma-type (grey), a coordination-game-type (black), or some other type of a situation (white). The data underlying the figure was generated with Python's numpy package (version 1.20; Harris et al., 2020). The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).

176x117mm (72 x 72 DPI)

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(a) **# of contributors** (rows) × **# of non-contributors** (columns)

# of contributors \ # of non-contributors	0	1	2	3	4	5	6	7	8
1	2.00	1.00	0.67	0.50	0.40	0.33	0.29	0.25	0.22
2	2.00	1.33	1.00	0.80	0.67	0.57	0.50	0.44	
3	2.00	1.50	1.20	1.00	0.86	0.75	0.67		
4	2.00	1.60	1.33	1.14	1.00	0.89			
5	2.00	1.67	1.43	1.25	1.11				
6	2.00	1.71	1.50	1.33					
7	2.00	1.75	1.56						
8	2.00	1.78							
9	2.00								

(b) **# of non-contributors** (rows) × **# of contributors** (columns)

# of non-contributors \ # of contributors	0	1	2	3	4	5	6	7	8
1	1.00	2.00	2.33	2.50	2.60	2.67	2.71	2.75	2.78
2	1.00	1.67	2.00	2.20	2.33	2.43	2.50	2.56	
3	1.00	1.50	1.80	2.00	2.14	2.25	2.33		
4	1.00	1.40	1.67	1.86	2.00	2.11			
5	1.00	1.33	1.57	1.75	1.89				
6	1.00	1.29	1.50	1.67					
7	1.00	1.25	1.44						
8	1.00	1.22							
9	1.00								

The benefit from groups with different combinations of agents from the perspective of (a) contributors and (b) non-contributors. The values in the tables are based on effort set to 1 and synergy from prosocial common-pool behavior set to 2.

171x168mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
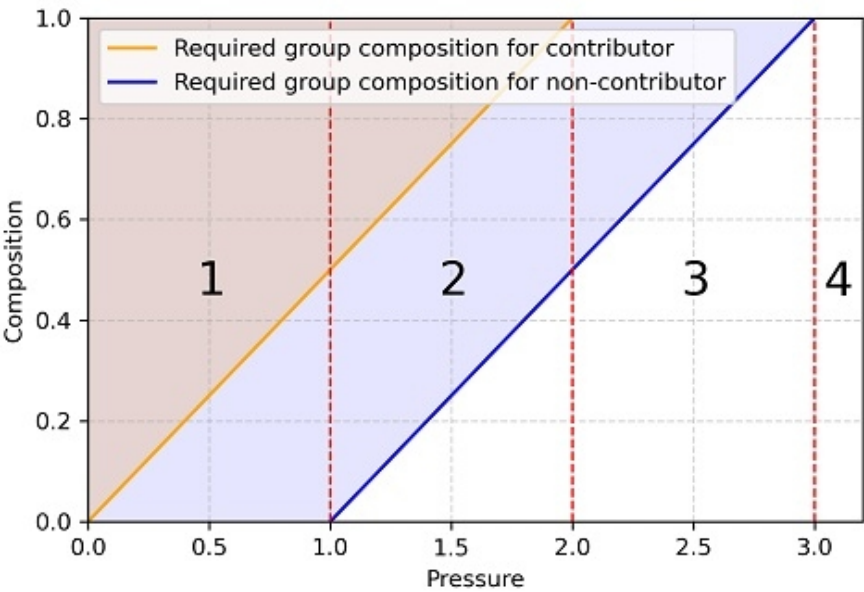47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3: The relationship between group composition and pressure to change behavior, from the perspective of contributors and non-contributors. The data underlying the figure was generated with Python's numpy package (version 1.20; Harris et al., 2020). The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).

142x95mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
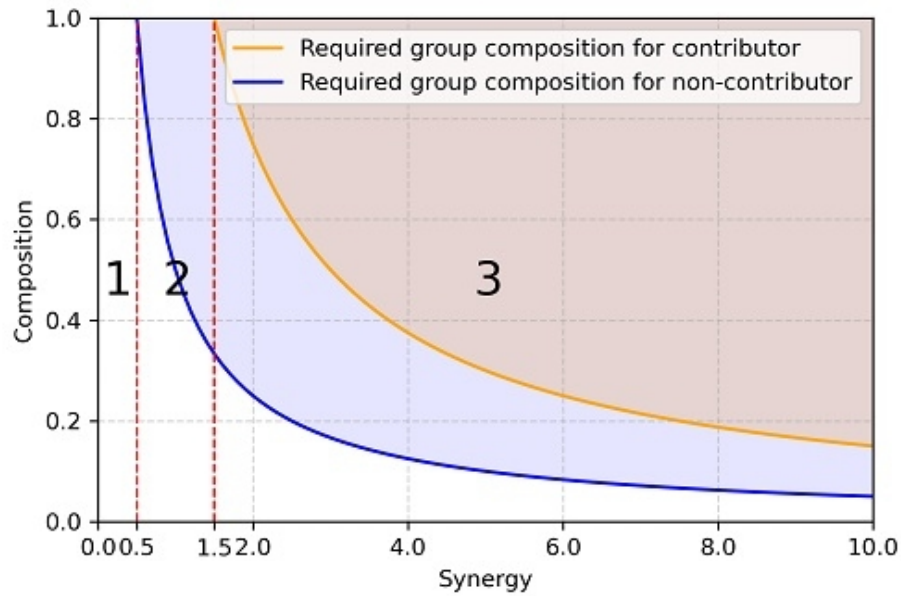50
51
52
53
54
55
56
57
58
59
60

Figure 4: The relationship between group composition and synergy from prosocial common-pool behavior, from the perspective of contributors and non-contributors. The data underlying the figure was generated with Python's numpy package (version 1.20; Harris et al., 2020). The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).

142x95mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
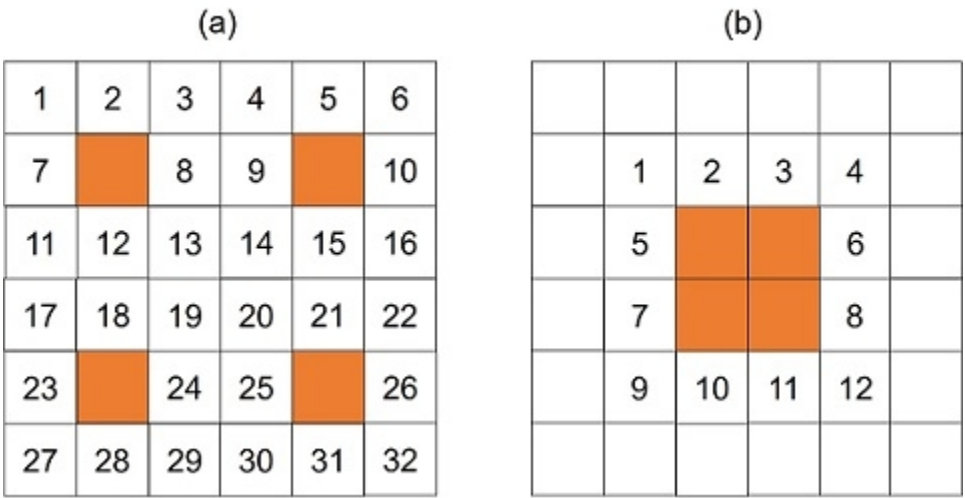45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 5: Two alternative group composition scenarios. Scenario (a) includes four spread out contributors (orange squares), offering 32 empty spots (8 each) next to them for other agents to fill. Scenario (b) includes four contributors clustered together - forming a square - offering only 12 empty spots next to them for others to fill. Scenario (a) benefits other agents, both contributors and non-contributors. Scenario (b) benefits the four contributors.
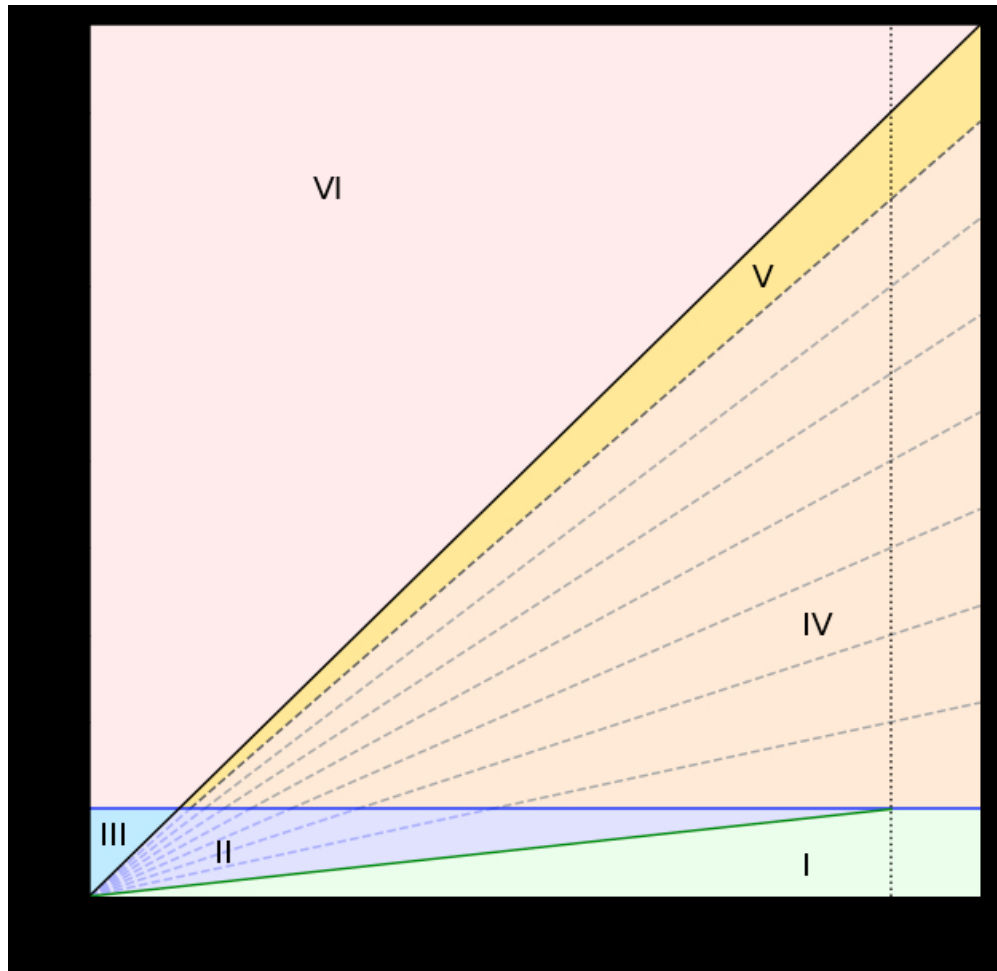
81x42mm (150 x 150 DPI)

Figure 6: Parameter space map of model equilibria by region (I-VI). The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).
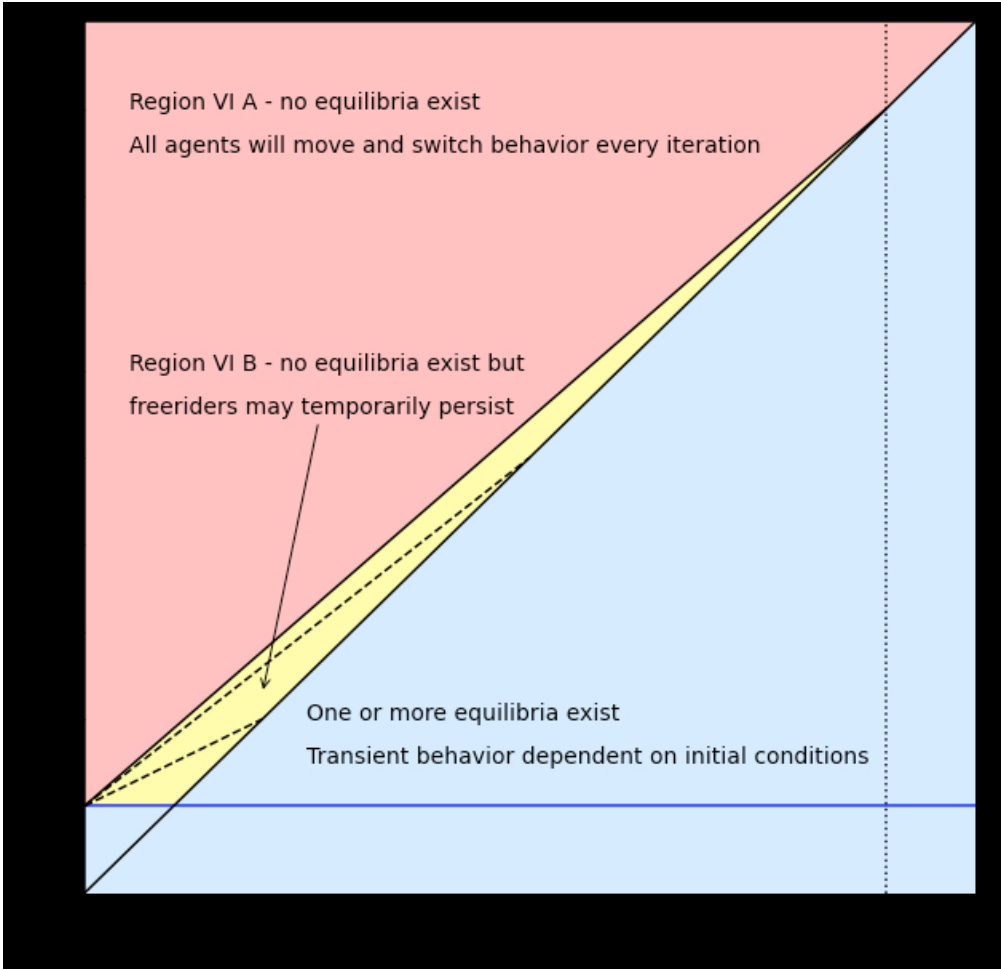
417x403mm (38 x 38 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 7: Parameter space map of model dynamics. The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).
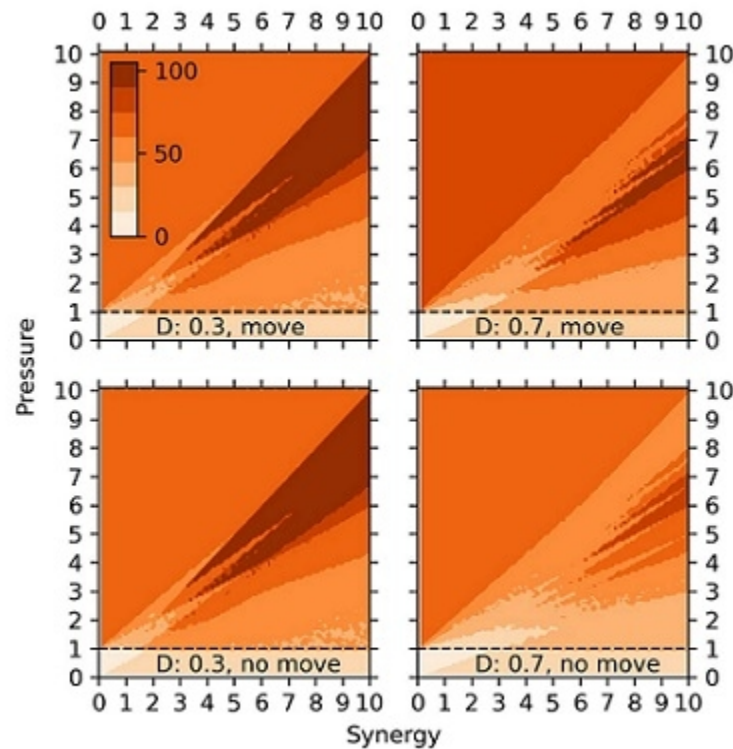
417x403mm (38 x 38 DPI)

Figure 8: Four contour maps of MGS I's parameter space of percent of contributions after 1,000 iterations. Each map illustrates the results (final percent of contributors) of 10,000 model runs under alternative pressure-synergy combinations (each map is actually based on 30,000 model runs, since each combination was simulated three times and averaged to account for stochasticity). The maps differ in density (D: 0.3, 0.7) and movement (move, no move) settings, with the following parameter settings kept constant: Z2 = 400, e = 1. The black horizontal dashed line depicts where effort equals pressure (e=p=1). The figure was generated with Python's matplotlib package (version 3.4.1; Hunter, 2007).

35x35mm (300 x 300 DPI)

1   **A new agent-based model offers insight into population-wide adoption of prosocial common-**

2   **pool behavior**

3   **Abstract:** Prosocial common-pool behavior is a subtractable and non-excludable resource that

4   benefits those exposed. While such behavior is ubiquitous, individuals who adopt it are usually at

5   a disadvantage to freeriders - those who benefit from it without adopting. We present a new

6   theoretical model - Multilevel Group Selection I - that establishes a relationship among important

7   factors influencing population-wide adoption of prosocial common-pool behavior (group

8   composition and size, movement between groups, multilevel group selection, population density,

9   pressure to change behavior, and synergy from prosocial common-pool behavior). The model

10 offers a surprisingly large degree of behavioral richness, given its simple structure and rules, and

11 serves as a fertile base for a fruitful and socially-relevant research program.

12 **Keywords:** agent-based model, common-pool resource, freeriding, multilevel group selection,

13 prosocial common-pool behavior, social change.

14

15   1 Introduction

16 We live in a rapidly changing world and adopting new behavior is vital for both individual and

17 collective benefit. Evidence shows, however, that many individuals at least initially do not adopt

18 new behavior (Juhola et al., 2016). There are a number of potential explanations for why. One is

19 that they may have a general skepticism or aversion to change (Rogers, 1983). Another is that they

20 may perceive the new behavior as a fad (Abrahamson, 1991) or a maladaptation (Barnett &

21 O'Neill, 2010; Magnan et al., 2016). Yet another explanation, which is the focus of this paper, is

22 that there is a cost to adopting the new behavior that they may want to avoid or at least delay

23 incurring. The size of this cost associated with adopting the new behavior is often influenced by

24 an individual's (intellectual, financial, physical) ability to provide the behavior, their need for it,

25 and their benefit from it.

26 The last explanation for why an individual might not adopt new behavior is especially probable

27 when the behavior is prosocial (i.e., behavior like helping, sharing, or cooperating that benefits

28 others [Twenge et al., 2007]). In such cases, adoption of the new behavior by others produces

29   positive externalities (i.e., positive side effects that benefit those exposed [Cornes & Sandler,

30   1996]), which provide exposed non-adopters with an opportunity and incentive to freeride (i.e.,

31   benefit from the behavior without adopting it [Battaglini et al., 2012; Kim & Walker, 1984; Ozono

32   et al., 2017]). Therefore, the ability to freeride likewise influences the associated cost and whether

33   an individual adopts. When the behavior is also a common-pool resource, i.e., is both subtractable

34   (the benefit from the behavior by one individual reduces possible benefit by another) and non-

35   excludable (the benefit from the behavior is not easily limited to specific individuals) (Ostrom et

36   al., 1994), freeriding by others increases the cost and reduces the benefit of the new behavior to

37   existing and potential adopters, in turn, further discouraging its adoption and reducing its presence

38   in a population.

39   To date, much of the theoretical progress in relation to prosocial common-pool behavior (e.g.,

40   cooperation [Axelrod, 1984]) has been made in relation to two individuals facing a situation with

41   a cost-benefit structure that is commonly referred to as the Prisoner's Dilemma (von Neumann &

42   Morgenstern, 1944). An example of this progress is the identification of the tit-for-tat strategy

43   (Axelrod & Hamilton, 1981; Dal Bó & Fréchette, 2019) that helps those implementing prosocial

44   common-pool behavior in such a situation avoid loss in interaction with those who are determined

45   to exploit them by not reciprocating. However, many (if not most) of the situations that individuals

46   face as they interact with others are not of the Prisoner's-Dilemma type (e.g., bargaining, coalition

47   building [Maskin, 2016]). Additionally, many groups that play important roles in society are

48   composed of more than two individuals, including (offline and online) communities and

49   classrooms, friendships, households, and professional organizations. Therefore, developing theory

50   of prosocial common-pool behavior for situations in groups of size greater than two is equally if

51   not more important than developing it for situations in groups of size two. However, simply

52   extending the Prisoner's Dilemma to groups of size greater than two by either coupling group

53   members into two sub-groups (e.g., Camerer, 2003; von Neumann & Morgenstern, 1944) or

54   engaging group members in pair-wise (as opposed to multi-wise) interaction (e.g., Helbing et al.,

55   2011; Helbing & Yu, 2008; Nowak & May, 1992; Nowak & Sigmund, 1998; Tarnita et al., 2009)

56   ignores the common-pool nature of such groups, i.e., it overlooks the effect group size and

57   composition have on a situation's cost-benefit structure (more on this in Sections 3.1 and 4.1.1).

58 Furthermore, many groups overlap as individuals hold membership in more than one group. Each
59 group provides its members with an opportunity to be a provider or a recipient of prosocial
60 common-pool behavior, as well as an opportunity to freeride or reciprocate. Experience from
61 interaction within any one of the groups also influences whether an individual repeats or adopts
62 the same or other prosocial common-pool behavior in that and/or other groups. The relative
63 proportion of members providing prosocial common-pool behavior in a group is also likely to
64 make the cost-benefit structure in some groups more favorable than in others (Boyd, 2018;
65 Goodnight & Stevens, 1997; Sober & Wilson, 1999; Waring et al., 2015; Wilson & Sober, 1994;
66 Wilson & Wilson, 2007). Such between-group interaction (e.g., Bowles & Gintis, 2009; Boyd &
67 Richerson, 1992; De Silva & Sigmund, 2009; Frank, 2009; Hauert et al., 2002, 2007; Wang et al.,
68 2011; Ye et al., 2011) is as important to a theory of prosocial common-pool behavior as the
69 aforementioned within-group interaction (exemplified by the Prisoner's Dilemma). Yet, current
70 theory on cooperative behavior does not address it (Maskin, 2016).

71 Therefore, there is a need to develop theory for prosocial common-pool behavior in groups of size
72 greater than two; under different levels of cost, benefit, and need; and taking into account
73 freeriding and differences in group composition. Such theory has the potential to help with
74 designing mechanisms aimed at preventing the exploitation of prosocial common-pool behavior
75 and facilitating its adoption. It would also have the potential to provide insight into the relationship
76 between prosocial common-pool behavior and tangible common-pool resources (more on this in
77 Sections 2 and 6).

78 To this end, we developed a new theoretical model in NetLogo (version 6.2.0; Wilensky, 1999),
79 called Multilevel Group Selection I (version 2.0; henceforth "MGS I" or "the base model") to
80 explore the influence of important factors on population-wide adoption of prosocial common-pool
81 behavior. Our long-term aim in developing the model was to create a fertile base upon which more
82 complex models - its variants - aimed at studying population-wide adoption of prosocial common-
83 pool behavior can be built. We then developed a variant of the model that sweeps through a preset
84 range of parameter combinations and used it to explore simulation outcomes.

85 In this paper, we first define prosocial common-pool behavior in Section 2. In Section 3, we
86 introduce the new MGS I model and describe our process of verifying it. In Section 4, we describe
87 our analysis of agent behavior and overall model dynamics. In Section 5, we present simulation

2

88 results and use them to verify the expected model dynamics and to analyze the influence of

89 population density, the initial percent of contributors, and movement. Finally, in Section 6, we

90 draw some initial conclusions from the analysis, discuss their relevance for regulating exploitation

91 of tangible common-pool resources, and share ideas for possible variants. The base model, its

92 variant (MGS I Sweep), and the Python (version 3.6; van Rossum & Drake Jr., 2009) scripts used

93 for analysis and visualization are available on GitHub: https://github.com/Multilevel-Group-

94 Selection.

2 The concept of prosocial common-pool behavior

96 We define prosocial common-pool behavior as a subtractable resource that benefits those exposed,

97 is provided by an individual through interaction with others, and the access to which is not easily

98 limited (i.e., is non-excludable). Prosocial common-pool behavior is both similar and different

99 from other types of common-pool resources, namely tangible (e.g., fisheries, forests, groundwater

100 basins [Ostrom, 1999]) and intangible (e.g., information [Hess & Ostrom, 2003]). For example,

101 both tangible common-pool resources and prosocial common-pool behavior can be depleted, i.e.,

102 they are both subtractable. However, in addition to being physically depletable (an individual

103 engaging in prosocial common-pool behavior can become physically or mentally exhausted), a

104 prosocial common-pool behavior can also be discouraged by social factors (such as who is the

105 depletor), new information, or changes in other (internal or external) factors. In other words, unlike

106 other types of common-pool resources, prosocial common-pool behavior has physical,

107 psychological, and social dimensions, and, therefore, has both tangible and intangible

108 characteristics. Prosocial common-pool behavior (subtractable and non-excludable) is also both

109 similar and different from other types of behavior, namely club (non-subtractable and excludable),

110 private (subtractable and excludable), and public (non-subtractable and non-excludable). For

111 example, access to both public and common-pool behavior cannot be easily limited to specific

112 individuals (i.e., they are both non-excludable). However, benefit from prosocial public behavior

113 is not subtractable (e.g., introducing laws, planting trees), while benefit from prosocial common-

114 pool behavior is (e.g., volunteering at a homeless shelter).

115 A clear example of prosocial common-pool behavior is collective voluntary labor, such as building

116 and cleaning (e.g., Simon & Mobekk, 2019). A less obvious example is prosocial competition

3

117   (e.g., Gilbert & Basran, 2019), which involves competing in a way that promotes a healthy

118   environment for others and oneself. Then, of course, there is adhering to shared rules in consuming

119   a tangible common-pool resource (e.g., Ostrom et al., 1994), which reduces the current depletion

120   of a resource, thereby benefiting others and oneself by making more of the resource available for

121   future consumption. Managing a tangible common-pool resource, therefore, requires, at least in

122   part, managing the depletion of the corresponding prosocial common-pool behavior (more on this

123   in Section 6.2).

124   Sources of motivation for prosocial common-pool behavior can vary substantially. Prosocial

125   common-pool behavior might be motivated by altruism, which is driven by concern for others;

126   egoism, which is driven by social status, reputation, or expectation for reciprocity; or by a desire

127   to adhere to a personal ideology. As mentioned above, adoption of new prosocial common-pool

128   behavior might be hindered by skepticism or aversion to change, perception of the new behavior

129   as a fad or maladaptation, or a desire to avoid or delay an associated cost. A prosocial common-

130   pool behavior can also provide immediate or future benefit to its provider, be either free to its

131   provider or at a cost, it can be synergistic, and it can have a local or global, and a current or future,

132   impact.

3 The Multilevel Group Selection I model

3.1 Description

135   Multilevel Group Selection I (version 2.0; MGS I) simulates a social space composed of agents

136   that either contribute prosocial common-pool behavior when in a group or not (i.e., freeride), and

137   that are equally exposed to some (physical, psychological, and/or social) pressure to change their

138   behavior. Exposure to prosocial common-pool behavior can help agents withstand the pressure,

139   especially when the behavior is synergistic. Agents move within their social space to avoid this

140   pressure, as a result forming into groups (common pools), which is where prosocial common-pool

141   behavior can occur. Agents unable to withstand the pressure automatically change from

142   contributing prosocial common-pool behavior to not or vice versa. We leave the empowerment of

143   agents with the ability to choose their behavior for future (more complex) variants of this base

144   model.

145 We conceptualize the social space that the agents move within as the maximum social structure

146 (network) of a population of agents. In our base model, this space takes the form of a square lattice,

147 Z^2, with each square representing a social spot (node), $z_{ij} \in Z^2$, that can be empty or occupied by

148 one agent. The lattice serves as a unique constraint on the structure of the population that for our

149 purposes conveniently focuses the study of within- and between-group (cluster) interactions on

150 groups ranging between 2 and 9 in size. We avoid boundary effects in the lattice by forming it into

151 a torus. The social space's density, D, determines how many of the space's spots are occupied by

152 agents during a simulation. Which spots are occupied is initially random and afterwards is

153 determined by agent movement. Exploring alternative topologies is an opportunity for future

154 variants.

155 There are N agents, labeled a_i, where $i = 1, 2, \ldots$ N and N is determined by the product of the

156 social space's density and size, $D*Z^2$. Initially, agents are randomly distributed (using a uniform

157 distribution) throughout the social space. Each agent's initial behavior is assigned randomly, using

158 the initial percent of contributors, $I_{t=0}$. At the beginning of each iteration, each agent is also

159 endowed with an equal level of effort, e_i, which in the base model we normalized across agents at

160 1. Agents need effort to provide prosocial common-pool behavior and withstand pressure. The

161 effort sets a limit to how much prosocial common-pool behavior an agent can contribute in a group

162 and there is a 1:1 relationship between the two. It takes all of an agent's effort to contribute

163 prosocial common-pool behavior to a group. On the other hand, non-contributors retain all of their

164 effort. We designed the base model and framed this study with humans in mind. However, we do

165 not see a reason why the base model or a variant could not apply equally well in the context of

166 other social life forms.

167 Groups represent all possible interactions among agents, which, in the real world, can vary

168 substantially in purpose and many other characteristics (e.g., duration, norms). A group is defined

169 from the perspective of its focal agent. A focal agent's group, $g_{i,t}$, at any point during a simulation

170 is composed of its Moore neighborhood neighbors and, as agents move, can vary from 2 to 9 agents

171 in size (including the focal agent) during a simulation. The dynamic number of groups at any point

172 during a simulation, $g_{i,t} \in G_t$, equals the number of agents who have at least one neighbor, i.e., G_t

173 $\leq N = D * Z^2$. Any agent that has a group is also in the groups of its neighbors (group members).

174 In other words, the groups of neighboring agents overlap and, as a result, an agent can be in 1 to 8

5

175   other groups (in addition to its own), depending on whether it has 1 or 8 neighbors, respectively

176   (Figure 1). An agent contributes (or not) an equal amount of effort in each one of its groups. An

177   agent can also be outside of a group, i.e., alone, in a social spot with no other agents within its

178   Moore neighborhood. An agent that is not in a group cannot contribute or benefit from prosocial

179   common-pool behavior.

180   **[Figure 1]**

181   An agent is able to withstand the pressure to change its behavior when the sum of its retained effort

182   and the benefit from its group is more than the pressure level. An agent can only derive benefit

183   from its own group, i.e., the group within which it is the focal agent. There are two leading

184   approaches to calculating the benefit from a group of a size greater than two. One is based on

185   sequential pairwise interaction between an agent and its neighbors, from which the individual

186   benefits to the agent are summed (e.g., Helbing et al., 2011; Helbing & Yu, 2008; Nowak & May,

187   1992; Nowak & Sigmund, 1998; Tarnita et al., 2009). This approach, however, does not account

188   for the subtractable and non-excludable characteristics of an agent's prosocial common-pool

189   behavior within a group. It does not account for the fact that the amount of effort available to an

190   individual agent within a group decreases with every interaction with another agent within that

191   group nor the fact that an agent's behavior within a group cannot be limited to one member at a

192   time. Within the context of common-pool behavior, the approach inflates both how much benefit

193   an agent can provide to and can receive from its group members, which biases agent behavior and

194   potentially simulation results in favor of larger groups and prosocial agents.

195   The other leading approach to calculating benefit from groups of size greater than two is based on

196   simultaneous multi-wise interaction among the agent and its neighbors, from which the individual

197   benefit to the agent is averaged (e.g., Bowles & Gintis, 2009; Boyd & Richerson, 1992; Choi &

198   Robertson, 2019; De Silva & Sigmund, 2009; Frank, 2009; Hauert et al., 2002, 2007; Wang et al.,

199   2011; Ye et al., 2011). This approach, which is a key building block in our base model, reflects

200   the subtractable and non-excludable characteristics of prosocial common-pool behavior within a

201   group. The resulting ability of a contributor and non-contributor to withstand pressure to change

202   its behavior are calculated as follows:

203                $ability\ of\ contributor = s * (\sum_{j=1}^{N_i} e_j) / N_i$ Equation 1

$$ability\ of\ non-contributor = e_i + s * (\textstyle\sum_{j=1}^{N_i} e_j)\ /\ N_i \qquad \text{Equation 2}$$

205 where e_i is the effort that each agent is endowed with at the beginning of each iteration and which,

206 when in a group, a contributor contributes and a non-contributor keeps; s is the level of synergy

207 from the prosocial common-pool behavior, and e_j is the effort contributed by a group member

208 (including the focal agent, if it is a contributor). The ability of an agent that is not in a group to

209 withstand the pressure to change its behavior is equal to the level of its unused effort, e_i, with

210 which it was endowed at the beginning of the iteration and which it did not have an opportunity to

211 contribute. As a result, the ability to withstand the pressure to change behavior is the same for all

212 agents that are alone, but may differ for agents in groups, based on their type and group's

213 composition. The amount of effort in possession resets for every agent each iteration, i.e., it does

214 not accumulate.

215 An iteration during simulation of the base model consists of two main steps, potential movement

216 and potential behavioral change. Agents can move within their social space, which represents any

217 process of leaving and/or joining a group. It could represent physical movement from one

218 geographic location to another or virtual movement from one online location to another. Agents

219 are selected randomly one-by-one and automatically move to one of the closest empty spots if their

220 ability to withstand pressure (the amount of effort in their possession) is below the pressure level.

221 As is the case with changing behavior, we leave the empowerment of agents with the ability to

222 choose where they move for future (more complex) variants of this base model. Having agents

223 move to the closest empty spot is intended to reflect limitations of local knowledge and could be

224 relaxed in future variants. After potentially moving, agents are selected randomly one-by-one and

225 automatically change their behavior if their ability to withstand pressure (the amount of effort in

226 their possession) is below the pressure level.

3.2 Verification

227

228 We took steps to verify that the base model correctly represented the conceptual model that we

229 described in Section 3.1. The model verification involved the following three sets of activities: (a)

230 code walkthroughs, and analyses of agent behavior under alternative (b) parameter settings and (c)

231 module deactivations. The code walkthroughs involved discussions of every code line's fit with

7

232 the corresponding components of and often led to simplifying changes to earlier versions of the

233 conceptual model. The analyses of agent behavior were conducted with the model set in slow-

234 mode and involved observing whether randomly selected agents behaved as was expected. The

235 analysis under alternative parameter settings involved varying parameters over their full plausible

236 ranges and observing results that were consistent with design objectives and analytically predicted

237 end states. Module deactivations included suppressing movement and/or behavior change, and

238 assuring the correctness of the results.

4 Theoretical analyses

240 We analyzed MGS I from a micro-level (agent behavior) and a macro-level (model dynamics)

241 perspective. Despite the base model's relative simplicity (in comparison to the complexity of real-

242 world social contexts and models that attempt to capture them), its outcomes vary substantially

243 under alternative pressure-synergy combinations and, as we show herein, provide plausible

244 explanations for real-world outcomes/interpretations.

4.1 Micro-level analysis of agent behavior

4.1.1 Link to game theory

247 Agents in MGS I can face situations that are commonly studied in game theory. As mentioned in

248 the introduction, much of the theoretical progress in relation to prosocial common-pool behavior

249 has been made in relation to two individuals facing a situation with a payoff structure that is

250 commonly referred to as the Prisoner's Dilemma (Table 1).

251   **[Table 1]**

252 To demonstrate that agents in MGS I can face the Prisoner's Dilemma, we first extended the

253 interpretation of the classic 2x2 payoff structure to groups of size greater than two. An agent's

254 ability to withstand pressure to change its behavior (i.e., its payoff, in game theoretic terminology)

255 equals: T (temptation) when it does not contribute prosocial common-pool behavior while at least

256 one agent in its group does, R (reward) when it and all the other agents in its group contribute, P

257 (punishment) when it and the other agents in its group do not contribute, and S (sucker's) when

8

258 the agent contributes while at least one agent in its group does not. This extension is applicable to

259 groups of all sizes.

260 We then identified the combinations of (group) size and synergy (from prosocial common-pool

261 behavior) in which agents can face the Prisoner's Dilemma (Figure 2). We did the same for another

262 (albeit less) popular situation in game theory, commonly referred to as the coordination game and

263 defined by the following payoff structure: R > T and P > S (Figure 2). Interestingly, outside of the

264 two extreme situations in which every group member either contributes or doesn't (and receives

265 an R or P, respectively), how many of the group members contribute (the group's composition)

266 does not affect whether they face a Prisoner's Dilemma (under the aforementioned size-synergy

267 combinations). Contribution decisions do, however, affect whether or not agents face a

268 coordination game situation, due to the resulting values of T and S.

269 **[Figure 2]**

270 In addition to potentially facing the Prisoner's-Dilemma or the coordination-game-type of

271 situations, MGS I agents can also find themselves in other types of situations. In fact, MGS I agents

272 face neither of the two situations in more than half of the size-synergy combinations (Figure 2),

273 suggesting a surprisingly large degree of behavioral richness in the model, despite its simple

274 structure and rules.

4.1.2 The influence of group composition

276 Group composition has a substantial impact on an agent's ability to withstand the need to change

277 its behavior and, in turn, on population-wide adoption of prosocial common-pool behavior. Table

278 2 illustrates this by listing the benefit to contributors and non-contributors from groups of different

279 compositions. If an agent's benefit from a group is equal to or less than the pressure to change its

280 behavior, the agent will need to change it.

281 **[Table 2]**

282 We used Equations 1 and 2 (which calculate an agent's ability to withstand pressure) to explore

283 the relationship between group composition and pressure. We first defined a group's composition

284 as the number of its contributing members divided by the total number of its members, $(\sum_{j=1}^{N_i} e_j)$ /

9

285   N_i. We then set the equations equal to pressure, since this is the point that determines if an agent

286   needs to change its behavior. Finally, we expressed the equations from the perspective of group

287   composition and through the pressure and synergy parameters:

288                 $Group\ composition\ of\ contributor = p/s$ Equation 3

289         $Group\ composition\ of\ non-contributor = (p - e)/s$ Equation 4

290   Figure 3 illustrates the relationships expressed in Equations 3 and 4 at different levels of group

291   composition and pressure to change behavior, and with the synergy from prosocial common-pool

292   behavior set equal to 2. A contributor or non-contributor that faces a situation in which their

293   composition-pressure combination is to the right of the orange or blue line, respectively, will need

294   to change its behavior.

295   **[Figure 3]**

296   As the broken red lines in Figure 3 indicate, composition-pressure combinations faced by agents

297   can be partitioned into four segments, separated by three critical pressure points.

298   ● The <u>1st segment</u> consists of composition-pressure combinations in which pressure is less than

299   effort ($p < e = 1$). This is the satisfaction boundary for lone agents. A contributor might not

300   be able to withstand pressure to change behavior if it faces a composition-pressure

301   combination that falls to the right of the orange line. A non-contributor facing a combination

302   from this segment does not need to change behavior.

303   ● The <u>1st critical point</u> is where pressure equals effort ($p = e = 1$). At this point and at higher

304   pressure, non-contributors need prosocial group members to withstand pressure to change

305   behavior.

306   ● The <u>2nd segment</u> consists of composition-pressure combinations in which pressure is equal to

307   or greater than effort but less than synergy ($e \leq p < s$). Both a contributor and a non-

308   contributor might not be able to withstand pressure to change behavior if it faces a

309   composition-pressure combination that falls to the right of the orange or blue line,

310   respectively.

311   ● The <u>2nd critical point</u> is where pressure equals synergy ($p = s = 2$). This is the satisfaction

312   boundary for contributing agents in groups with 100% prosocial behavior. At this point and at

313   higher pressure, contributors are not able to withstand pressure to change behavior.

10

314   ●   The <u>3rd segment</u> consists of composition-pressure combinations in which pressure is equal to

315   or greater than effort but less than the sum of effort and synergy from prosocial common-pool

316   behavior ($e \leq p < e + s$). All contributors are not able to withstand pressure to change

317   behavior, while a non-contributor is not able to if the combination it faces falls to the right of

318   the blue line.

319   ●   The <u>3rd critical point</u> is where pressure equals the sum of effort and synergy ($p = e + s$). At

320   this point and at higher pressure, both contributors and non-contributors are not able to

321   withstand pressure to change behavior.

322   ●   The <u>4th segment</u> consists of composition-pressure combinations in which pressure is equal to

323   or greater than the sum of effort and synergy ($e + s \leq p$). All contributors and non-

324   contributors are not able to withstand pressure to change behavior.

325 Ceteris paribus, a change (increase/decrease) in the effort level shifts the blue line (right/left). It

326 affects the non-contributors directly by increasing/decreasing their safe zone, and both contributors

327 and non-contributors indirectly through the benefit they receive from their groups. An increase

328 (decrease) in the synergy from prosocial common-pool behavior decreases (increases) the slope of

329 the orange and blue line, thereby increasing the safe zones and the size of segments 2 and 3.

330 Similarly, Figure 4 illustrates the relationships expressed in Equations 3 and 4 with different group

331 composition and synergy from prosocial common-pool behavior combinations, and this time with

332 pressure (instead of synergy) set equal to 1.5. As in Figure 3, a contributor or non-contributor that

333 faces a situation in which their composition-synergy combination is to the left of the orange or

334 blue line, respectively, will need to change its behavior.

335   **[Figure 4]**

336 As the broken red lines in Figure 4 indicate, composition-synergy combinations faced by agents

337 can be partitioned into three segments, separated by two critical value points.

338   ●   The <u>1st segment</u> consists of composition-synergy combinations in which synergy is lower than

339   the difference between pressure and effort, $s < p - e$. In such cases, both contributors and non-

340   contributors are better off avoiding groups.

341   ●   The <u>1st critical point</u> is equal to the difference between pressure and effort, $p - e$. At this point

342   and at higher synergy, it becomes beneficial to not-contribute.

11

343 ● The <u>2nd segment</u> consists of composition-synergy combinations in which synergy is equal to

344 or greater than the difference between pressure and effort, but is less than pressure,

345 $p - e \leq s < p$. In such cases, non-contributors begin to benefit from groups with large numbers

346 of contributors. On the other hand, contributors remain better off avoiding groups.

347 ● The <u>2nd critical point</u> is where synergy equals pressure, $s = p$. At higher synergy levels, it

348 becomes beneficial for contributors to join groups.

349 ● The <u>3rd segment</u> consists of composition-synergy combinations in which synergy is greater

350 than pressure, $p < s$. In such cases, both contributors and non-contributors benefit from groups

351 with large numbers of contributors.

352 Ceteris paribus, a change (increase/decrease) in effort shifts (right/left) the first critical point. It

353 affects the non-contributors directly by expanding/reducing their safe zone, and both contributors

354 and non-contributors indirectly through the benefit they receive from their groups. An increase

355 (decrease) in pressure reduces (increases) the concavity of the orange and blue lines and shifts

356 them leftward, thereby reducing the safe zones.

357 Finally, in addition to influencing the cost-benefit structure of its group members, a group's

358 composition also influences the options available to non-group members. As mentioned before,

359 the larger the number of contributors in an agent's group, the greater the agent's ability to

360 withstand the pressure to change behavior. This makes empty spots near contributors highly

361 valuable. The number of such spots changes substantially with the composition of existing groups

362 (Figure 5), influencing the options available to non-contributors to freeride and the options

363 available to contributor's and non-contributor's to improve their ability to withstand pressure. This

364 highlights an important tradeoff that exists at lower levels of contributors between the ability of an

365 average agent to withstand pressure and that of contributors that have managed to group with other

366 contributors.

367 **[Figure 5]**

4.2 Macro-level analysis of model dynamics

4.2.1 Existence and classification of equilibria

The model is at equilibrium when all agents are satisfied with their group affiliation and behavior choice. If we assume homogeneous effort across all agents (prosocial or not), then both Equations 1 and 2 are linear in effort and without loss of generality this coefficient can be scaled to one. The value of prosocial common-pool behavior is then a linear function of the synergy and the proportion of the group adopting prosocial behavior. A freeriding group member receives this value plus the value of their own effort. In groups with no prosocial agents, the focal member only receives the value of their own effort, which is identical to the case of a lone agent without a group. The value of the prosocial common-pool behavior is therefore the entire potential benefit of group membership.

For the model to be at equilibrium, either there are no prosocial agents and the pressure is less than or equal to the individual effort (extinction), or there exist agents with prosocial behavior and across all groups the minimum value of the prosocial common-pool behavior is greater than or equal to the pressure (persistence or full adoption). Based on these observations, we can characterize the existence of different equilibria across the synergy-pressure parameter space (Figure 6).

[Figure 6]

Key elements in Figure 6 include: (a) the satisfaction boundary for lone agents and groups without prosocial behavior, $p = 1$; (b) the satisfaction boundary for contributing agents in groups with 100% prosocial behavior, $p = s$; and (c) the synergy level above which all agents will be better off adopting prosocial common-pool behavior regardless of the composition of their group, $s = 9$. We can break up the possible equilibria into six general regions.

Regions **I** through **III** are all below the satisfaction boundary for non-contributing agents, so no non-contributing agent will move or switch behavior in these regions. In region **I**, all behavior choices and group compositions produce sufficient reward to keep agents from moving or changing behavior. Pressure is too low to produce dynamics. All initial agent distributions are at equilibrium (persistence of prosocial common-pool behavior).

13

396   Region **II** has enough pressure to cause contributing agents without sufficient prosocial
397   participation in their group to move and change behavior. Non-contributing agents are always
398   satisfied whether or not their groups provide them benefit as freeriders. Dynamics from typical
399   random initial distributions will exhibit short transients during which contributing agents move
400   and change behavior. Most observed equilibria contain few contributing agents in stable groups
401   (persistence of prosocial common-pool behavior). The required proportion of contributing agents
402   for a stable group increases with pressure.

403   Region **III** is above and to the left of the $p = s$ limit for contributing agents to be at equilibrium,
404   so every contributing agent will move and change behavior. All initial distributions lead to 100%
405   non-contributing behavior after one iteration (extinction of prosocial common-pool behavior).

406   Regions **IV** through **VI** are all above the satisfaction boundary for non-contributing agents, which
407   implies they cannot withstand the pressure to change without a freerider benefit in these regions.
408   As region **VI** is also above and to the left of the $p = s$ limit for contributing agents to be at
409   equilibrium, no equilibria exist in this region (persistence of prosocial common-pool behavior).

410   Region **IV** admits the existence of stable groups with contributing agents and freeriding non-
411   contributing agents. The degree of freeriding permissible without destabilizing a group depends
412   on the pressure. Dynamics from typical random initial distributions may exhibit short or long
413   transients before reaching an equilibrium or the model dynamics may lead to the formation of
414   stable groups in a configuration where one or more agents are trapped in a confined social space
415   with no accessible satisfactory behavior and group combination available. In region **V** only 100%
416   prosocial groups are stable (full adoption of common-pool behavior). There is no opportunity for
417   freeriding at equilibrium.

418   4.2.2 Characterization of dynamics in absence of equilibria

419   Model dynamics in region **VI** fall in to two distinct categories, dynamics in which every agent
420   moves and switches behavior on every iteration due to pressure exceeding the maximum common-
421   pool benefit a freerider can receive (Region **VI A** in Figure 7), and dynamics in which freeriding
422   agents can sustain their behavior for an iteration due to participation in a sufficiently prosocial
423   group (Region **VI B** in Figure 7). The sufficiently prosocial groups that permit freeriders in region

14

424 **VI B** cannot persist as their members with prosocial behavior are unable to resist the pressure to

425 move and change behavior.

[Figure 7]

427 The ultimate boundary for region **VI B** is given by the maximum benefit for a freerider,

428 $1 + (8/9) * s$. Obtaining this payoff requires a maximal sized group of contributing agents.

429 Smaller groups may temporarily support freeriders when the pressure is closer to the region **VI**

430 lower boundary as illustrated by the dashed lines in region **VI B**, which represent the payoff

431 boundaries for groups with 80% and 50% contributing agents.

5 Simulation results

433 MGS I's structure allows for the following three alternative and substantially different outcomes:

434 1. When one or more equilibria are present, dynamics from typical random initial distributions

435 may exhibit short or long transients before reaching an equilibrium (regions **I-V**).

436 2. In some regions, even though one or more equilibria exist, transients may lead to the formation

437 of stable groups in a configuration where one or more agents are trapped in a confined social

438 space with no accessible satisfactory behavior and group combination available, which leads

439 to a persistent random cycle through a subregion of state space containing no equilibria (region

440 **IV**).

441 3. When the pressure is high enough that no equilibria exist (region **VI**), all agents change their

442 behavior either every iteration period, or some freeriders persist for one or more iterations

443 before switching behaviors in a perpetual random walk.

5.1 Sweeping through the parameter space

445 To verify the theoretical analysis in Section 4 and further explore the relationship between

446 simulation outcomes and specific parameter-setting combinations, we developed a variant of MGS

447 I, called MGS I Sweep (version 1.0), which sweeps through a range of pressure-synergy

448 combinations and, to account for stochasticity, does each sweep a preset number of times (which

449 we set to 3) and then averages the results. We used MGS I Sweep to sweep through pressure-

450 synergy combinations ranging from 0-0 to 10-10, at 0.1 increments, and at two alternative density

15

451   (0.3, 0.7) and movement (move, no move) settings (Figure 8). We implemented the two movement

452   settings by deactivating movement and re-sweeping through the same ranges of pressure-synergy

453   combinations.

454   **[Figure 8]**

455   The dynamics that lead to the outcomes that are visualized by the contour maps (Figures 8) are

456   explained by the regions described in Section 4.2. The areas where equilibria are reliably reached

457   include region **III** (extinction of prosocial common-pool behavior), which is clearly visible near

458   the origin. Regions **I** and **II** also reach equilibria (persistence of prosocial common-pool behavior)

459   but are not well differentiated by looking at the final percent of contributors due to the percent of

460   contributors in the initial state. Some loss of contributors from the initial state is visible in that part

461   of region **II** closest to the $s = p$ boundary.

462   Regions **IV** and **V** show where model dynamics lead to population-wide adoption of prosocial

463   common-pool behavior and where the dynamics lead to partial adoption (Case 1 above) or partial

464   adoption with some agents trapped in untenable social spaces (Case 2). Note that 100% prosocial

465   behavior is always an equilibrium state in regions **IV** and **V**. The contours in regions **IV** and **V** can

466   be described using the stability boundaries of various sized prosocial groups and freerider payoffs,

467   and represent both types of dynamics.

468   Finally, the third case of no equilibrium is easily spotted the upper left regions in all 4 maps

469   corresponding to region **VI A** of our analysis. Region **VI B** is noticeable on both low-density maps

470   (left top and bottom) as the area between region **VI A** and the $s = p$ boundary, and though less

471   obvious, is also present in the high density maps.

472   **5.2 The influence of population density**

473   The influence of population density on social life is multidimensional and complex (e.g., Ahlfeldt

474   & Pietrostefani, 2019), and can be simultaneously positive (increased innovation, reduced per

475   capita pollution, access to amenities) and negative (higher congestion, and prices of land and

476   housing) (Duranton & Puga, 2020). Its influence on prosocial common-pool behavior, however,

477   is not widely (if at all) studied. In the model, population density influences the conditions that

478   agents face during a simulation in two important ways. One is it influences the number of agents

                                                                                              16

479   being simulated. This is because the number of agents is determined by the product of density and

480   the number of spots in the social space ($D*Z^2$). An increase in density increases the chance of

481   having a neighbor, which, from the perspective of an agent, has both positive and negative side

482   effects. The positive side effect is the corresponding increase in the chance of having a contributor

483   as a neighbor. The negative side effect is the increase in the chance of a neighbor being a non-

484   contributor. The second way in which density influences the conditions that agents face is through

485   the number of possible moves during a simulation (which density reduces) and thereby the chance

486   of finding an empty spot near a contributor or non-contributor (for better or for worse,

487   respectively).

488   The influence of population density on population-wide adoption of prosocial common-pool

489   behavior is illustrated by the differences in the contour maps that were generated under the two

490   alternative density settings (0.3, 0.7) (Figure 8). Overall, increases in density reduce the number

491   of pressure-synergy combinations that lead to population-wide adoption of prosocial common-

492   pool behavior. One likely cause is that higher population density levels increase the chance of

493   having a non-contributing (i.e., a freeriding) neighbor, which reduces the ability of their neighbors

494   to withstand pressure to change. Because within any group a contributor is at a disadvantage to a

495   non-contributor, this effect has a greater negative impact on the former than the latter.

5.3 The influence of the initial percent of contributors

497   The relative numbers of socially or culturally different individuals in a population are critical in

498   shaping their interaction and, in turn, future outcomes (Kanter, 1977). In the model, the influence

499   of the initial percent of contributors on population-wide adoption of prosocial common-pool

500   behavior affects underlying behavior and the final outcomes, but not the size of the regions or the

501   transition zones between them (as the case with density). Final outcomes remain the same in

502   regions **III** and **V**, in which the final percent of contributors reach 0% and 100%, respectively

503   when equilibrium is reached. In region **I**, the final percent of contributors is simply equal to the

504   initial percent as all initial conditions are at equilibrium. Equilibria obtained for region **II** are

505   dependent on initial conditions as the initial percent is an upper limit on the final percent.

506   In region **VI A**, the initial percent of contributors influences the position of the two equilibria

507   points between which the percent of contributors fluctuates during a simulation. Specifically, the

17

508   locations of the equilibria points are symmetrically distanced from 50%. For example, if the initial

509   percent of contributors is 40% (or 60%), the percent of contributors during a simulation fluctuates

510   between 40% and 60% (or vice versa). If the initial percent is 30% (or 70%), the percent fluctuates

511   between 30% and 70% (or vice versa). In all cases, the average final percent of contributors is

512   50%. However, depending on the time period at which a simulation ends, the final percent of

513   contributors can be equal to either one of the equilibria. For example, the simulations that Figure

514   8 relies on ran for 1,000 time periods. With the simulations starting on an odd number (1) and

515   ending on an even number (1,000), the final percent of contributors are equal to the other

516   equilibrium point, e.g., if the initial percent of contributors is 30%, the final percent of contributors

517   is 70%, and vice versa.

518   When interpreting these identical and symmetric fluctuations in population-wide behavior, it is

519   important to keep in mind that we are describing the base model, based on which more realistic

520   variants can be built. For example, a variant that assigns a probability to whether an agent changes

521   its behavior would remove the identical and symmetric nature of the described fluctuations in the

522   percent of contributors. Aside from their identical and symmetric nature, however, population-

523   wide shifts in behavior as demonstrated by MGS I are anything but absent from human history.

524   Primary examples include cultural and political revolutions (Kuran, 1995; Opp & Gern, 1993),

525   which are increasingly "scaled up" by social media (Mundt et al., 2018). The aforementioned

526   equilibria can be seen as "tipping points" (Centola et al., 2018) that lead to major shifts in social

527   convention.

5.4 The influence of movement

529   Movement in MGS I represents any process of leaving and/or joining a group. Agents are selected

530   randomly one-by-one and automatically move to one of the closest empty spots if their ability to

531   withstand pressure is below the pressure level. Movement between groups is not required for

532   population-wide adoption. Even random movement can increase adoption of prosocial common-

533   pool behavior. This is evident by the *all-around* lower final percent of contributors in the bottom

534   row of Figure 8 contour plots, which are based on simulations without movement, when compared

535   to those generated with the same density levels and initial percent of contributors, but with

536   movement in the top row. This result challenges the recently proposed notion that greedy (i.e.,

18

537   goal-oriented, non-random) movement is necessary for increases in adoption of prosocial

538   common-pool behavior (e.g., Helbing et al., 2011).

6 Discussion

539

6.1 Persistence and adoption of prosocial common-pool behavior

540

541   Individuals providing prosocial common-pool behavior are often at a substantial disadvantage to

542   those who exploit it. Yet, prosocial common-pool behavior persists in most populations and new

543   examples appear and are often adopted. MGS I offers a plausible explanation for why this occurs.

544   It demonstrates how at sufficiently high levels of synergy from prosocial common-pool behavior,

545   groups with larger numbers of contributors are better able to withstand pressure than groups with

546   lower numbers. This mechanism is commonly-referred to in biology as multilevel group selection

547   (Boyd, 2018; Goodnight & Stevens, 1997; Sober & Wilson, 1999; Waring et al., 2015; Wilson &

548   Sober, 1994; Wilson & Wilson, 2007). Outside of a few exceptions (e.g., Boyd, 2018; Waring et

549   al., 2015), models developed for studying multilevel group selection are focused on non-human

550   populations. Therefore, one of our main contributions is a new base model for studying multilevel

551   group selection in human populations.

552   The base model establishes a previously unexplored relationship among important factors

553   influencing prosocial common-pool behavior, including group composition, group size, movement

554   between groups, multilevel group selection, population density, pressure to change behavior, and

555   the synergy from prosocial common-pool behavior. Our theoretical analysis of agent behavior

556   (Section 4.1) described the influence: (a) size and synergy have on the types of situations

557   contributors and non-contributors of prosocial common-pool behavior might face, (b) composition

558   and synergy have on the ability of the contributors and non-contributors to withstand pressure, and

559   (c) composition and pressure have on the benefit of joining a group.

560   Our theoretical analysis of the model's dynamics (Section 4.2) identified six distinct regions in the

561   model's parameter space, in which pressure-synergy combinations lead to substantially different

562   outcomes: extinction (region **III**), persistence (regions **I & II**), and full adoption (regions **IV & V**)

563   of a prosocial common-pool behavior. Regions **IV & V** demonstrate the presence in the model of

564   multilevel group selection: under sufficiently high synergy from prosocial common-pool behavior,

19

565   adopters organize into sufficiently large groups that permit them to offset the disadvantage they

566   have against non-adopters in withstanding pressure to change their behavior and lead to its

567   population-wide adoption.

568   Our analysis of the model's simulation results (Section 5) verified the theoretical analysis in

569   Section 4.2 and provided additional insight into the influence of the aforementioned factors on

570   population-wide adoption of prosocial common-pool behavior. It demonstrated that: (a) increases

571   in density reduce the number of pressure-synergy combinations that lead to population-wide

572   adoption of prosocial common-pool behavior, (b) the initial percent of contributors affects the

573   underlying behavior and the final outcomes, but not the size of the regions or the transition zones

574   between them, and (c) random (non-greedy) movement assists adoption of prosocial common-pool

575   behavior. Overall, the analyses demonstrate that further exploration of the relationships among the

576   important factors has the potential to be of great value in efforts to promote prosocial common-

577   pool behavior.

6.2 Relevance for managing tangible common-pool resources

579   The MGS I model also has the potential to provide insight into the management of common-pool

580   resources (e.g., fisheries, forests, groundwater basins). The benefit derived from consuming a

581   common-pool resource is private to the consumer, while the cost associated with its consumption

582   is shared by all consumers (Hardin, 1968). For this reason, management efforts aimed at preventing

583   exploitation of a tangible common-pool resource tend to fall into one or a combination of the

584   following two categories: (a) efforts that limit an individual's consumption and (b) those that force

585   an individual to internalize the associated cost.

586   However, outside of a perfect market economy, the decision to consume a tangible common-pool

587   resource is rarely ever solely based on the aforementioned cost-benefit structure alone. Resisting

588   overconsuming is a prosocial common-pool behavior that, as we have demonstrated in Section 4.1,

589   is encouraged/discouraged by: (a) the broader synergy from the prosocial common-pool behavior

590   to the individual and (b) the pressure to change their behavior. Both include the benefit and cost

591   associated directly with the consumption of a resource. As mentioned in Section 2, prosocial

592   common-pool behavior might additionally and, perhaps, in some cases, be entirely motivated by

593   altruism, egoism, or by a desire to adhere to a personal ideology. As demonstrated in Section 5,

20

594 social structure, movement, and multilevel group selection also influence the prevalence of
595 prosocial common-pool behavior.

596 Therefore, populations aiming to prevent exploitation of tangible common-pool resources have
597 additional psychological and social levers that complement those related to limiting consumption
598 of the resource and internalizing its cost. Many social levers are already incorporated in formal
599 and informal norms. Examples include social punishment for overconsumption and social reward
600 for restraint. Many psychological levers are also already ingrained during upbringing (e.g., Simon
601 & Mobekk, 2019) and in some cases are also genetically transmitted through prosocial genes
602 (Knafo-Noam et al., 2015). The analyses in Sections 4.2 and 5 suggest that the right pressure-
603 synergy combinations can lead populations to higher levels of adoption of prosocial common-pool
604 behavior, and, through association, responsible consumption of tangible common-pool resources.
605 It additionally suggests that social structure plays a considerable role in encouraging/discouraging
606 adoption.

6.3 Potential variants

608 MGS I has a simple and modular structure that eases the process of developing its variants that
609 replace existing processes, add complexity to them, and/or introduce new processes while
610 maintaining important components in situations with prosocial common-pool behavior, i.e.,
611 density, movement, multilevel group selection, pressure, and synergy. We experimented with a
612 number of what would now be considered as variants in the process of designing MGS I, including
613 ones that introduce: greedy (i.e., goal-oriented, non-random) movement, probability in behavioral
614 change, and limitations of group access to non-contributors. Each one of these added a layer of
615 real-world complexity that influenced model results, but were excluded from MGS I in pursuit of
616 a fundamental design.

617 Here, we briefly mention two possible directions for future variants not mentioned above. The first
618 direction involves variants that empower MGS I agents with an ability to assess the likely behavior
619 of group members and choose their own group behavior accordingly. These variants would permit
620 the study of game theoretic situations in a population consisting of multiple groups of varying size.
621 The new processes would replace the existing process of behavioral change, which automatically
622 changes an agent's behavior when pressure is equal to or above its ability to withstand it. To study

21

623 the Prisoner's Dilemma or the coordination game, specifically, a variant would additionally need

624 to limit the ranges of the pressure and synergy parameters, as well as introduce a mechanism for

625 regulating group size (see Section 4.1.1). A resulting variant could then be used to explore games

626 in multi-group populations under alternative pressure-synergy combinations, and density and

627 initial percent of contributors settings.

628 The second direction involves variants that empower MGS I agents with a wider range of

629 cognitive, social, demographic, and potentially additional behavioral processes. These variants

630 would permit the study of the influence of these processes on adoption of prosocial common-pool

631 behavior. One option could be to replace the base model's agents with SOSIEL agents (Sotnik,

632 2018), each of which makes decisions using a cognitive architecture composed of nine processes

633 (anticipatory learning, goal prioritizing, counterfactual thinking, innovating, social learning, goal

634 selecting, satisficing, signaling, and individual or collective action taking) and can be born, die,

635 and pair with other agents. This second direction could of course be configured to also align with

636 the first direction, permitting the study of the influence of cognitive, behavioral, social, and

637 demographic processes on adoption of prosocial common-pool behavior in game theoretic

638 situations. These aforementioned directions for future variant development are only two among

639 many, serving as examples of how MGS I provides a fertile base for a fruitful and socially-relevant

640 research program.

641 **References**

642 Abrahamson, E. (1991). Managerial Fads and Fashions: The Diffusion and Rejection of

643   Innovations. *The Academy of Management Review*, *16*(3), 586–612.

644 Ahlfeldt, G. M., & Pietrostefani, E. (2019). The economic effects of density: A synthesis.

645   *Journal of Urban Economics*, *111*, 93–107. https://doi.org/10.1016/j.jue.2019.04.006

646 Axelrod, R. M. (1984). *The Evolution of Cooperation*. Basic Books.

647 Axelrod, R. M., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science, New Series*,

648   *211*(4489), 1390–1396.

649 Barnett, J., & O'Neill, S. (2010). Maladaptation. *Global Environmental Change*, *20*(2), 211–213.

22

650   https://doi.org/10.1016/j.gloenvcha.2009.11.004

651 Battaglini, M., Nunnari, S., & Palfrey, T. (2012). *The Free Rider Problem: A Dynamic Analysis*

652   (Working Paper No. 17926; Issue 17926). National Bureau of Economic Research.

653   http://www.nber.org/papers/w17926

654 Bowles, S., & Gintis, H. (2009). Beyond Enlightened Self-Interest: Social Norms, Other-

655   Regarding Preferences, and Cooperative Behavior. In S. A. Levin (Ed.), *Games, Groups,*

656   *and the Global Good* (pp. 57–78). Springer Berlin Heidelberg.

657   https://doi.org/10.1007/978-3-540-85436-4

658 Boyd, R. (2018). *A different kind of animal: How culture transformed our species*. Princeton

659   University Press.

660 Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or

661   anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171–195.

662   https://doi.org/10.1016/0162-3095(92)90032-Y

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

664   University Press.

665 Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). Experimental evidence for

666   tipping points in social convention. *Science*, *360*(6393), 1116–1119.

667   https://doi.org/10.1126/science.aas8827

668 Choi, T., & Robertson, P. J. (2019). Contributors and Free-Riders in Collaborative Governance:

669   A Computational Exploration of Social Motivation and Its Effects. *Journal of Public*

670   *Administration Research and Theory*, *29*(3), 394–413.

671   https://doi.org/10.1093/jopart/muy068

672 Cornes, R., & Sandler, T. (1996). *The Theory of Externalities, Public Goods, and Club Goods*.

23

673   Cambridge University Press.

674   Dal Bó, P., & Fréchette, G. R. (2019). Strategy Choice in the Infinitely Repeated Prisoner's

675   Dilemma. *American Economic Review*, *109*(11), 3929–3952.

676   https://doi.org/10.1257/aer.20181480

677   De Silva, H., & Sigmund, K. (2009). Public Good Games with Incentives: The Role of

678   Reputation. In S. A. Levin (Ed.), *Games, Groups, and the Global Good* (pp. 85–103).

679   Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-85436-4

680   Duranton, G., & Puga, D. (2020). The Economics of Urban Density. *Journal of Economic*

681   *Perspectives*, *34*(3), 3–26. https://doi.org/10.1257/jep.34.3.3

682   Frank, S. A. (2009). Evolutionary Foundations of Cooperation and Group Cohesion. In S. A.

683   Levin (Ed.), *Games, Groups, and the Global Good* (pp. 3–40). Springer Berlin

684   Heidelberg. https://doi.org/10.1007/978-3-540-85436-4

685   Gilbert, P., & Basran, J. (2019). The Evolution of Prosocial and Antisocial Competitive Behavior

686   and the Emergence of Prosocial and Antisocial Leadership Styles. *Frontiers in*

687   *Psychology*, *10*, 610. https://doi.org/10.3389/fpsyg.2019.00610

688   Goodnight, C. J., & Stevens, L. (1997). Experimental Studies of Group Selection: What Do They

689   Tell US About Group Selection in Nature? *The American Naturalist*, *150*(S1), S59–S79.

690   https://doi.org/10.1086/286050

691   Hardin, G. (1968). The Tragedy of the Commons. *Science*, *162*(3859), 1243–1248.

692   https://doi.org/10.1126/science.162.3859.1243

693   Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D.,

694   Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van

695   Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., …

24

696   Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362.

697   https://doi.org/10.1038/s41586-020-2649-2

698   Hauert, C., De Monte, S., Hofbauer, J., & Sigmund, K. (2002). Replicator Dynamics for

699   Optional Public Good Games. *Journal of Theoretical Biology*, *218*(2), 187–194.

700   https://doi.org/10.1006/jtbi.2002.3067

701   Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., & Sigmund, K. (2007). Via Freedom to

702   Coercion: The Emergence of Costly Punishment. *Science*, *316*(5833), 1905–1907.

703   https://doi.org/10.1126/science.1141588

704   Helbing, D., & Yu, W. (2008). Migration as a mechanism to promote cooperation. *Advances in*

705   *Complex Systems*, *11*(04), 641–652. https://doi.org/10.1142/S0219525908001866

706   Helbing, D., Yu, W., & Rauhut, H. (2011). Self-Organization and Emergence in Social Systems:

707   Modeling the Coevolution of Social Environments and Cooperative Behavior. *The*

708   *Journal of Mathematical Sociology*, *35*(1–3), 177–208.

709   https://doi.org/10.1080/0022250X.2010.532258

710   Hess, C., & Ostrom, E. (2003). Ideas, Artifacts, and Facilities: Information as a Common-Pool

711   Resource. *Law and Contemporary Problems*, *66*(1/2), 111–145.

712   Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science &*

713   *Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

714   Juhola, S., Glaas, E., Linnér, B.-O., & Neset, T.-S. (2016). Redefining maladaptation.

715   *Environmental Science & Policy*, *55*, 135–140.

716   https://doi.org/10.1016/j.envsci.2015.09.014

717   Kanter, R. M. (1977). Some Effects of Proportions on Group Life: Skewed Sex Ratios and

718   Responses to Token Women. *American Journal of Sociology*, *82*(5), 965–990.

25

719   https://doi.org/10.1086/226425

720   Kim, O., & Walker, M. (1984). The free rider problem: Experimental evidence. *Public Choice*,

721   *43*(1), 3–24. https://doi.org/10.1007/BF00137902

722   Knafo-Noam, A., Uzefovsky, F., Israel, S., Davidov, M., & Zahn-Waxler, C. (2015). The

723   prosocial personality and its facets: Genetic and environmental architecture of mother-

724   reported behavior of 7-year-old twins. *Frontiers in Psychology*, *6*.

725   https://doi.org/10.3389/fpsyg.2015.00112

726   Kuran, T. (1995). The Inevitability of Future Revolutionary Surprises. *The American Journal of*

727   *Sociology*, *100*(6), 1528–1551.

728   Magnan, A. K., Schipper, E. L. F., Burkett, M., Bharwani, S., Burton, I., Eriksen, S., Gemenne,

729   F., Schaar, J., & Ziervogel, G. (2016). Addressing the risk of maladaptation to climate

730   change: Addressing the risk of maladaptation to climate change. *Wiley Interdisciplinary*

731   *Reviews: Climate Change*, *7*(5), 646–665. https://doi.org/10.1002/wcc.409

732   Maskin, E. (2016). How Can Cooperative Game Theory Be Made More Relevant to

733   Economics? : An Open Problem. In J. F. Nash, & M. Th. Rassias (Eds.), *Open Problems*

734   *in Mathematics* (pp. 347–350). Springer International Publishing.

735   https://doi.org/10.1007/978-3-319-32162-2_9

736   Mundt, M., Ross, K., & Burnett, C. M. (2018). Scaling Social Movements Through Social

737   Media: The Case of Black Lives Matter. *Social Media + Society*, *4*(4),

738   205630511880791. https://doi.org/10.1177/2056305118807911

739   Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, *359*.

740   Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring.

741   *Nature*, *393*, 573–577.

26

742 Opp, K.-D., & Gern, C. (1993). Dissident Groups, Personal Networks, and Spontaneous

743   Cooperation: The East German Revolution of 1989. *American Sociological Review*,

744   *58*(5), 659. https://doi.org/10.2307/2096280

745 Ostrom, E. (1999). Revisiting the Commons: Local Lessons, Global Challenges. *Science*,

746   *284*(5412), 278–282. https://doi.org/10.1126/science.284.5412.278

747 Ostrom, E., Gardner, R., & Walker, J. (Eds.). (1994). *Rules, Games, and Common-Pool*

748   *Resources*. The University of Michigan Press.

749 Ozono, H., Kamijo, Y., & Shimizu, K. (2017). Punishing second-order free riders before first-

750   order free riders: The effect of pool punishment priority on cooperation. *Scientific*

751   *Reports*, *7*(1), 14379. https://doi.org/10.1038/s41598-017-13918-8

752 Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed). Free Press ; Collier Macmillan.

753 Simon, C., & Mobekk, H. (2019). Dugnad: A Fact and a Narrative of Norwegian Prosocial

754   Behavior. *Perspectives on Behavior Science*, *42*(4), 815–834.

755   https://doi.org/10.1007/s40614-019-00227-w

756 Sober, E., & Wilson, D. S. (1999). *Unto others: The evolution and psychology of unselfish*

757   *behavior*. Harvard University Press.

758 Sotnik, G. (2018). The SOSIEL Platform: Knowledge-based, cognitive, and multi-agent.

759   *Biologically Inspired Cognitive Architectures*, *26*, 103–117.

760   https://doi.org/10.1016/j.bica.2018.09.001

761 Tarnita, C. E., Antal, T., Ohtsuki, H., & Nowak, M. A. (2009). Evolutionary dynamics in set

762   structured populations. *Proceedings of the National Academy of Sciences*, *106*(21), 8601–

763   8604. https://doi.org/10.1073/pnas.0903019106

764 Twenge, J. M., Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Bartels, J. M. (2007). Social

27

765   exclusion decreases prosocial behavior. *Journal of Personality and Social Psychology*,

766   *92*(1), 56–66. https://doi.org/10.1037/0022-3514.92.1.56

767   van Rossum, G., & Drake Jr., F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

768   von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*.

769   Princeton University Press.

770   Wang, J., Wu, B., W. C. Ho, D., & Wang, L. (2011). Evolution of cooperation in multilevel

771   public goods games with community structures. *EPL (Europhysics Letters)*, *93*(5), 58001.

772   https://doi.org/10.1209/0295-5075/93/58001

773   Waring, T. M., Kline, M. A., Brooks, J. S., Goff, S. H., Gowdy, J., Janssen, M. A., Smaldino, P.

774   E., & Jacquet, J. (2015). A multilevel evolutionary framework for sustainability analysis.

775   *Ecology and Society*, *20*(2), art34. https://doi.org/10.5751/ES-07634-200234

776   Wilensky, U. (1999). *NetLogo* (6.2.0) [NetLogo]. Center for Connected Learning and Computer-

777   Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/

778   Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral

779   sciences. *Behavioral and Brain Sciences*, *17*(4), 585–608.

780   https://doi.org/10.1017/S0140525X00036104

781   Wilson, D. S., & Wilson, E. O. (2007). Rethinking the Theoretical Foundation of Sociobiology.

782   *The Quarterly Review of Biology*, *82*(4), 327–348. https://doi.org/10.1086/522809

783   Ye, H., Tan, F., Ding, M., Jia, Y., & Chen, Y. (2011). Sympathy and Punishment: Evolution of

784   Cooperation in Public Goods Game. *Journal of Artificial Societies and Social Simulation*,

785   *14*(4), 20. https://doi.org/10.18564/jasss.1805

786

28