Supplementary and summarized ODD-protocol

Following the template provided by Grimm et al. (2020)[1], we add an ODD file to:

- de Matos Fernandes, C. A., Flache, A., Bakker, D. M., & Dijkstra, J. (2021). A bad barrel spoils a good apple: How uncertainty and networks affect whether meritocratic matching can foster cooperation. *Working paper*.
- de Matos Fernandes, C. A. (2021). Two agent-based models of cooperation in dynamic groups and fixed social networks (version 1.0.0). *CoMSES Computational Model Library*. https://www.comses.net/codebase-release/66d84cb4-c45f-46bb-bef6-b9e298758fa5/

Model description

The model is extensively described in our paper (de Matos Fernandes et al., 2021) wherein we address the following: the model in general, some simulation examples, typology of agents, threshold model equations, decision procedure, implementation of reinforcement learning, description of time steps per iteration, an overview of the random spatial graph algorithm, and how we implement homophily in the network. Moreover, we describe the matching rules figuratively as well as in-text-wise. A full overview of the variables is provided in the paper. What is more, the reader can find comments per line in the model code (de Matos Fernandes, 2021).

The overall purpose of both models is to assess the impact of complete and incomplete information during meritocratic matching on cooperation levels of prosocial agents. If advancement towards more cooperative groups solely relies on merits that do not fully capture individual cooperative abilities, then, our model BehaviorSpace experimental findings suggest that harvesting information from homophilous social networks is one solution for prosocial agents to match accordingly and fully capture their cooperative potential. For this, two NetLogo ABMs are added to CoMSES. The single group ABM allows us to inspect the consequence of having a set number of prosocials and proselfs in the group. The bad barrels ABM allows us to test our key intuitions, described in de Matos Fernandes et al. (2021).

To consider our model realistic enough for its purpose, we use the following patterns in cooperation levels and prosociality segregation after meritocratic matching.

[1] Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., Deangelis, D. L., … Ayllón, D. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation, 23*(2). https://doi.org/10.18564/jasss.4259

The model includes the following entities: proself and prosocial individuals, groups, and network links. The state variables characterizing these entities are listed in **Table 1**. As for the spatial and temporal resolution and extent: A single tick represents several individual actions until all acted in the group. Therefore, a single time step results in a single action for the individual and group. Also, in the network, a single time step corresponds to a single action in the 2-person prisoner dilemma (if the dyad is selected).

The most important processes of the model, which are repeated every time step, are the behavioral and learning algorithm in the group and network. Agents first decide to cooperate or defect based on their previous action, associated payoff, and updated threshold. Following the individual action, agents adjust their threshold according to what others did in the group and via the learning algorithm. This means that prosocials – who have a lower threshold and therefore cooperate more easily – are more likely to cooperate at first, but even prosocial increase their threshold and defect more readily when their cooperative actions do not pay off. The same counts for proself agents with a higher threshold (defect more easily). If defection does not pay off, proselfs are more likely to lower their threshold and learn to cooperate. The new threshold is used for the subsequent tick. Another key process is meritocratic matching. Every *X* ticks, agents can try to leave and join another group. Especially information during meritocratic matching is important. We implemented six matching rules, ranging from full to incomplete information.

The most important design concept of the model is the way we combine information derived from groups and networks. Other important design features are: reinforcement learning, probabilistic threshold model decision-making, heterogeneity (prosocials and proselfs), and stochasticity in decision-making and meritocratic matching. For more information about key concepts, please see the model section in de Matos Fernandes et al. (2021).

Also, model dynamics are predominantly driven by meritocratic matching moments, generally allowing a spike of cooperation for prosocial agents. For instance, model out in the bad barrels ABM shows sudden increases in cooperation levels after the meritocratic matching moment. However, our findings also show that prosocials quickly fall in line with their fellow group mates and defect more than wanted.

Finally, please find an overview of key variables implemented in NetLogo and the paper in **Table 1**.

Table 1. Overview of key variables.

Entity	Variable NetLogo	Sign	Description	Value
Individual	students		Turtles are defined as students	
	threshold	$\tau_{i,t}$	Adaptive threshold	(0, 1)
	initial-threshold	τ_i	Fixed (initial) threshold	{0.3, 0.7}*
	prob-c	$p_{i,t}$	Probability to cooperate	(0, 1)
	cooperate?	$c_{i,t}$	Cooperation/defection decision group	{0, 1}
	coop-rate	C_{10}	Mean cooperation level, last 10 iterations	(0, 1)
	coop-rate-total	C_{all}	Mean cooperation level, overall	(0, 1)
	private-c?	$c_{i,t,sn}$	Cooperation/defection decision network	{0, 1}
	private-reputation	$C_{10,sn}$	Mean cooperation level, last 10 iterations, network	(0, 1)
	cooperativeness-score-all	GC_{10}	Mean cooperation from group and network, last 10 iterations	(0, 1)
	noise		Noise in leave-stay decision	{0, 0.01, 0.05, 0.25}
	length-last-rounds-history		Memory of agents	10
	L	l	Learning rate	{0.1, 0.3, 0.5, 0.7, 0.9}
	M	m	Slope (noise in decision algorithm)	{1, 5, 10}
	COST	h	Cost of cooperation (input net share)	3
	BENEFIT-OF-D	d	Payoff when all defect (input net share)	-0.5
	BENEFIT	b	Benefit of cooperation (input net share)	4.5
	my-group		Group label	{0, 1, …, 19}
Group and individual	average-group-rate-lastXrounds; group-reputation	G_{10}	Mean cooperation level in the group, last 10 iterations	(0, 1)
	count-members; number-of-group-members		Count of group members	8
Group	studyties		Group membership connections	560
	number-of-students; number-of-agents-per-group	FS	Fixed group size	8
	how-many-other-c		Count of cooperators in the group	{0, 1, …, 8}
	participation-rate	k_t; $G_{i,t}$	Mean cooperation level in the group, given time point	(0, 1)
	num-of-groups	G	Count of groups in the population	20

	prosociality-segregation	MS_{group}	MS index to calculate the chances for a same-type tie in the group context	$0 < MS_{group}$
	sorted-groups		Preference list of groups available. Ordered based on average level of cooperation in the group.	{highest, …, lowest}
Population	count-ticks-for-selection	X	Moment of matching	{100, 200, 300}
	max-number-of-ticks	E	End iteration of a single run	400
	%-prosocials	PA	Proportion prosocial agents (1 – PA is the proportion proself agents)	{0.2, 0.4, 0.6, 0.8}
	selection?		Matching rules	{1, 2, …, 6}
	tick	t	A single iteration	
	leave-stay		decision rule why agents decide to stay or leave their group	{t_i, 1-t_i, 0.5}
	N-prosocials		Number of prosocials in a single group	{0, 1, …, 8}
Network	socialties		Social ties for the social network	800
	r	r	Chance of network dyad selection	{0.01, 0.05, 0.25, 0.5}
	structural-homophily?		If 'on', then prosocials are scattered on a pre-defined area on the grid	on; off
	range-sn		Information range about network cooperation of others in the network	2
	behavioral-homophily?		If 'on' then agents only interact with agents who behaved previously similarly	on; off
	k	k	Tied to k other agents	5
	y	w	Distance to others indicator. Higher values stress the importance of geographically nearby agents for tie creation.	8
	moody-segregation	$MS_{network}$	MS index to calculate the chances for a same-type tie in the social network	$0 < MS_{network}$

* The first value refers to prosocials (τ_i = 0.3) and the second to proselfs (τ_i = 0.7).