### Citations

Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF: „The ODD protocol: a review and first update“

Volker Grimm, Steven F. Railsback, Christian E. Vincenot, Uta Berger, Cara Gallagher, Donald L. DeAngelis, Bruce Edmonds, Jiaqi Ge, Jarl Giske, Jürgen Groeneveld, Alice S.A. Johnston, Alexander Milles, Jacob Nabe-Nielsen, J. Gareth Polhill, Viktoriia Radchuk, Marie-Sophie Rohwäder, Richard A. Stillman, Jan C. Thiele and Daniel Ayllón.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) 7 <https://www.jasss.org/23/2/7.html> DOI: 10.18564/jasss.4259

### ODD Template

The model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006). And the updated version (Grimm et al. 2020)

# 1. Purpose and patterns

The purpose of the model is to explore the conditions and mechanisms that lead to global consensus across identity labels, and varying degrees of within-label consensus and between-label diversity. We start with a population where choices are uniformly distributed in the population and agents are randomly assigned a label. Hence, labels and choices are initially uncorrelated.

Societal patterns: individuals who identify with specific identity labels or subcultures (e.g., musicians, scientists) tend to share common choices and attitudes more than those who do not share the same label. Factors assumed to explain this pattern are (a) Shared experiences: they create a common ground for individuals to bond and develop a sense of belonging; this may lead to the adoption of common choices and attitudes within the group. As part of social learning, these repeated interactions produce a shared sampling of similar social information, leading these individuals to conform. (b) Socialization and peer influence: people tend to interact with like-minded peers and these interactions tend to reinforce and validate their choices and attitudes. (c) Values and beliefs: shared values and beliefs can shape the choices and attitudes of individuals within the subculture leading to similar choices and attitudes on topics related to art, fashion, and social norms. (d) Symbolic representation: Subcultures often have distinct symbols, fashion styles, music genres, and other forms of cultural expression. These symbols act as markers of identity and facilitate identification among members. By adopting and displaying these symbols, individuals signal their affiliation with a particular subculture and align

themselves with the associated choices and attitudes. In this model, we focus specifically on symbolic representation as the primary driver of identity formation and choice alignment.

It's important to note that while individuals within an identity label often share common choices, there can still be diversity in their choices and attitudes. Nevertheless, the shared identity label often results in a greater likelihood of common choices and attitudes among its members.

# 2. Entities, state variables, and scales

Entities represent individuals and each individual is characterized by a set of issues, corresponding choices, and a label.

ISSUES AND CHOICES: Individuals have choices about a set of issues “num-issues”. Issues vary from 1 to 10 and for each issue they have an specific choice (e.g., A or B), varying from 2 to 10 choices. An issue represents a problem agents face, and choices represent the potential solution or attitude towards the specific problem. For each issue, the agent holds a belief in which choice is the best one, i.e., the preferred choice. This belief is a mental-model or bayesian representation of the probability that each of the choices is the best one. For instance, for an issue with choice A and B, an agent holds a numerical vector with the probability values for choice A and B, e.g., [1, 2] respectively. In this particular case, the agent holds the belief that choice B is the best one. Agents hold a numerical vector per issue, where the values in the vector represent the probabilities of each choice.

LABELS: agents hold a numerical variable representing a label, this variable ranges from -1 to *number-of-labels* in the system. Where a value of -1 represents no label and a value ≥ 0 a numerical label. Labels are not fixed and can change over time.

LABEL COLLECTIVES: Depending on the initial settings of initialization, interactions between agents holding the same label have a different outcome than interactions between agents holding different labels or no labels. When the process *selective-processing* is on and interacting agents hold the same label, the sender has an influence *strength-of-influence* times stronger over the receiver than in an interaction where the sender and receiver hold different labels. Further, when the *multi-issue-discourse* process is on, agents holding the same label will discuss more than one issue per interaction. This number is set by the parameter *number-of-issues-to-discuss*. Finally, when the process *Ignoring* is on, receivers will ignore the information from senders that hold a different label. In other words, interactions between agents of different labels have no effect.

Every interaction, each entity selects another at random to interact. Thus, every entity can potentially interact with every other entity in the model. In this sense, the model represents a fully connected network.

The model is run for an indefinite number of time-steps, until stability of the output measures is observed. Time steps in the model do not correspond to a fixed real-world temporal scale. Each time step represents a cycle in which every agent engages in one interaction, exchanging

information and updating their choices. The interpretation of a time step depends on the assumed frequency of meaningful interactions in real-world decision-making. If we assume that agents interact approximately once per day, then 1000 time steps would roughly correspond to ~3 years. However, if meaningful interactions occur less frequently (e.g., weekly discussions), the same 1000 time steps could represent a much longer period. The model remains agnostic to a specific time scale, allowing flexibility in its interpretation based on the context of opinion formation.

# 3. Process overview and scheduling

Time is modelled in discrete steps. Every time-step, the following processes occur in the given order.

1) Agents without a label have a probability (*prob-adopting-label)* of adopting a label.
2) Each agent randomly selects a partner for an interaction. Depending on whether agents hold a label or not, interactions and updates of variables may occur in three different ways:
    a) Both agents have no label, or the receiver has a label, but the sender has no label: a simple interaction occurs. The sender agent selects one issue at random, and for this issue, it selects the choice with the highest strength. The sender then displays this choice to the receiver. Upon receiving the displayed choice, the receiver increases the strength of that specific choice in its own records by 1.
    b) Receiver has no label, sender has a label. The receiver agent adopts the label of the sender with probability *prob-adopting-label*. If the receiver adopts the label of the sender, an interaction between agents holding the same label occurs (see c); otherwise, a simple interaction occurs (see a).
    c) Receiver and sender have a label. Here there are two different scenarios:
        i) Receiver and sender hold the same label. If selective processing is on, the sender selects one issue at random and for this issue, it selects the choice with the highest strength. The sender displays this choice to the receiver. The receiver increases the strength of that specific choice in its own records by *strength-of-influence-same-label*. If multiple-discourse is on, agents discuss issues *number-of-issues* times. Each time, the sender selects one issue at random and for this issue it selects the choice with the highest strength and displays it to the receiver. The receiver increases the strength of that specific choice in its own records by *strength-of-influence-same-label*.
        ii) Receiver and sender hold different labels. If *ignoring* is on, no interaction occurs. If *ignoring* is off, a simple interaction occurs (see a).
3) Agents with a label have a probability (*Prob_dropping_Label)* of dropping their label.

# 4. Design concepts

***Basic principles.*** The model's underlying design revolves around the theories of affective polarization and democratic discourse. According to the theory, an emotional attachment towards the in-group may result in a preference towards the in-group's choices and hostility towards the outgroup's choices. This may lead to the following outcomes: a) to an alignment of differences: "identities, beliefs, and cultural preferences become drawn into an all-encompassing societal division." among different groups (Törnberg, 2022); (b) to an homogenization of choices within groups which leads to less information diversity within the group; and (c) to a breakdown of political discourse across groups, restrained to opinion in-groups and a decreased ability to come to consensus.

Further, according to Albarracin et al., 2022, labels in a society act as signals of distinction badges of in-group membership (e.g., churches, sport clubs, regional identities, etc.). Labels may serve as indicators of a shared worldview, influencing choice formation and group identity. In a society, labels can affect choice formation and affective polarization. They can do so optimally by maintaining some degree of diversity across sub-groups, avoiding locks-in in choices, and allowing for social influence both within and across groups. In this case, societies cooperate in the negotiation of choices. In a suboptimal way on the other hand, labels may lead to either a society where there is total conformity, i.e., and information poor society or to a society with lock ins, where choices cannot be revised or aligned, interactions and social influence happens only within the labels, leading thus to communities demarcated by their beliefs and antagonistic towards the beliefs of others.

Informational conformity is a potential mechanism underlying the social dynamics in opinion formation (REF). Information conformity refers to the need that individuals have to cohere with the beliefs of one's in-group. This need is driven by the desire to obtain accurate information about reality and occurs when individuals are uncertain and rely on trusted members of a group to infer what is true, what action to take, or what opinion to have. When presented with information, individuals found this information to be more convincing if it is given by the in-group. Thus, the ingroup's information is generally trusted and has a higher influence than that presented by outgroup members (REF).

Our model takes all these theories into account. In the model, agents hold a certain number of issues for which they have a mental representation of what is the best choice. Individuals interact by exchanging information on what they consider to be the best choice on a given issue. However, the amount of information exchange with others as well as the degree of influence that they have towards others depend on the labeling. Where interactions among individuals with the same label have a higher influence and/or exchange of information than interactions among individuals with different labels or with no label.

With this model, we will investigate the conditions and mechanisms leading to three different societal scenarios. (a) societies where all individuals converge on the same choice for all given issues, i.e., total conformity. (b) societies where individuals within labels conform to an specific choice for all issues but where labels differ in the best choices for all issues, i.e., societies with

labels demarcated by their own beliefs. (c) societies where both labels and society vary in the best choice for all issues, i.e., societies with great diversity of choices.

The model is theoretical and not calibrated to any specific society, thus its use in real-world scenarios is questionable. Nevertheless, the model can shed light on the basic principles governing opinion formation. The model is an extension of the continuous opinions discrete actions (CODA) model. The CODA model has been extensively studied under different settings and conditions (REF). The model allows the emergence of different final stages for a society of agents. In a not strongly connected network, CODA would always lead to disagreement, while, on a fully connected network, consensus would always happen in the long run.

***Emergence.*** As mentioned previously, we expect the emergence of three different opinion variation scenarios at the societal level. (a) societies where all individuals converge on the same choice for all given issues, i.e., total conformity. (b) societies where individuals within labels conform to an specific choice for all issues but where labels differ in the best choices for all issues, i.e., societies with labels demarcated by their own beliefs. (c) societies where both labels and society vary in the best choice for all issues, i.e., societies with great diversity of choices.

The emergence of these scenarios will be dependent mainly on the initial values of some parameters: *strength-of-influence-same-label, number-of-issues* discussed among same label individuals, *prob-dropping-label*, etc. However, the relationships among these variables and how they derive into different societal scenarios are unknown and will be studied.

***Adaptation.*** Individuals hold a number of issues where each issue has a specific number of choices. For each issue agents have a preferred choice. Preferred choices of agents change as they hold interactions with others. The degree of change as well as the amount of information exchange between agents depend on whether agents hold the same label of different ones (or no label). Agents are more influenced and/or exchange more information with those sharing the same label. Preferred choices have no impact on individual success, in fact, in this model there is no *fitness* advantage in holding an specific choice for any given issue.

***Objectives.*** As previously mentioned, preferred choices do not have any impact on individuals' success. In fact, there is no measurement of success in the model. Our main objective is to study the evolution of preferred choices at the societal and label level given the rules of interaction and influence set at initialization.

***Learning.*** Agents change their preferred choices as they exchange information with other agents. So, agents learn the preferred choices of others and based on these experiences they modified their preferred choice.

***Prediction.*** Agents do not make predictions. They only observed the preferred choices of *others.*

***Sensing*.** When interacting, receivers perceived the preferred choices of the senders. Further, the receiver agent perceives whether the sender belongs or not to a certain label and whether the label of the receiver is the same one as the one they belong.

***Interaction*.** Interactions occur in dyads of individuals, where one is the sender (the one displaying a preferred choice) and the other the receiver (the one sensing the preferred choice of the sender). Interactions are considered direct because by sharing its preferred choice, the sender may influence the preferred choice of the receiver.

***Stochasticity*.** Individuals choose interacting partners at random. Further, the adoption and disposal of labels are partially stochastic processes since they are governed by a probability and thus occur with a specific frequency. The adoption of a new label outside the context of an interaction is done at random.

***Collectives*.** Interactions among individuals sharing the same label have a higher influence and/or exchange of information than interactions among individuals with different or no label. Thus, individuals with the same label belong to a group/community that affect each other more strongly. The number of labels in the system from which an individual can choose are given at initialization, thus labels are not emergent. However, individuals can adopt or dispose of labels with a probability set at initialization.

***Observation*.**

We are focusing on two key aspects: the coherence of choices within groups sharing the same label and the diversity across different labels.

To assess, how similar are agents holding the same label, for each group (label) and each issue we take frequencies of choices and calculate Shannon-Weiner diversity index, which is basically entropy normalized by logarithm of number of possible choices, so it ranges between [0, 1]. A value of 0 indicates unanimous agreement within a label on a given issue, while a value of 1 suggests a uniform distribution of choices. For the sake of clarity in our visual representations, we present the Shannon-Wiener diversity index averaged over issues and labels.

To measure the likeness between choices of two distinct groups with different labels on a specific issue, we examine the distribution of agent support for each possible choice within both groups. Then we calculate cosine similarity between resultant vectors representing frequencies of choices. A similarity score of 1 denotes identical vectors, while 0 means that they are orthogonal. Subsequently, we compute the average similarity across issues and all possible pairwise comparisons.

# 5. Initialization

At initialization the following parameters are set and fixed throughout the simulation: (1) population size (constant = 2500); (2) number of issues [1, 5, 10], (3) number of choices per

issue [2, 5, 10], (4) number of labels [2, 6, 10], (5) multi-issue discourse [1, 5, 20]; (6) strength of influence [1, 5, 20]; (7) ignoring (on/off); (8) probability of dropping label [0.000, 0.0001, 0.0005, 0.0010, 0.0050, 0.0100]; (9) probability of adopting label (constant = 0.1).

Once these parameters are set, the agents are initialized. For each agent, the strength of choices per issue are set by drawing values from a uniform distribution [0-3], and each agent is assigned a label at random.

# 6. Input data

The model does not rely on empirical input data and is entirely synthetic.

# 7. Sub models

There are no sub models in the present model.

---

[1] References are given in the manuscript.

[2] This is a reference to the ODD “first update” manuscript.