Hapres | Journal of Sustainability Research
sustainability.hapres.com

Article

Sustainable Altruism: Simulating the Future of Prosociality and Peacebuilding—A Conceptual Review

F. LeRon Shults [1,2,3,*], Justin E. Lane [3,4], Michael Gantley [1,3], Katherine O'Lone [5]

[1] Institute for Global Development and Social Planning, University of Agder, Universitetsveien 25, Kristiansand 4630, Norway; mgantley@culturepulse.ai (MG)

[2] NORCE Center for Modeling Social Systems, Universitetsveien 19, Kristiansand 4630, Norway

[3] CulturePulse, Inc., 131 Continental Drive, Suite 305, Newark, DE 19713, United States; justin@culturepulse.ai (JEL)

[4] Slovak Academy of Science, Institute of Ethnology and Social Anthropology, Štefánikova 898/49, Staré Mesto 811 04, Slovakia

[5] School of Social Sciences, University of Manchester, Oxford Road, Manchester, M13 9PL, UK; katherine.olone@manchester.ac.uk (KO)

* Correspondence: F. LeRon Shults (FLS), Email: leron.shults@uia.no.

Abstract

In this article, we argue for the potential value of participatory multi-agent artificial intelligence (MAAI) modeling for addressing the pragmatic challenge of promoting a sustainable altruism among individuals and groups that will enable us to find pathways of collective action that lead toward more peaceful coexistence. Along the way, we note that this approach to modeling also addresses the scientific challenge within computer science and social simulation of creating more psychologically realistic artificial agents whose interactions occur in more realistic social networks. We identify some of the evolved cognitive and coalitional biases that make it so difficult to achieve an equilibrium of sustainable altruism in contemporary human societies, describe some of the innovative ways in which recent advances in MAAI approaches to psychological and cultural modeling open up new opportunities for simulating solutions to these challenges, and address some of the ethical issues associated with using modeling and simulating methods to help us proactively navigate the Anthropocene.

KEYWORDS: sustainability; altruism; prosociality; social simulation; Anthropocene; multi-agent artificial intelligence; computer modeling; ethics

Open Access

Received: 03 April 2025
Accepted: 15 May 2025
Published: 21 May 2025

INTRODUCTION

Recent developments in computational modeling and social simulation (CMSS) methodologies have enormous potential for contributing to our understanding of the conditions under which—and the mechanisms by which—complex socio-ecological systems can transform in sustainable ways. Such approaches can also provide policy-relevant insights for some of the major challenges facing humanity in the Anthropocene, the current era in which human behavior has become a major factor in determining macro-trends in the global ecosystem, including exacerbated climate changes that threaten the environmental systems on which our own species depends in order to thrive—and perhaps to survive [1,2]. Calls for new ways of thinking and acting to facilitate our species' adaptation during (and hopefully through) the Anthropocene are growing in number and urgency [3–5]. CMSS techniques are the most powerful set of explanatory tools we currently have for analyzing and predicting complex adaptive systems with nonlinear dynamics and causal feedback loops, and so it is not surprising that attempts have been made to apply them to the challenge of human adaptation in this new and rapidly shifting environment [6–9].

However, there are at least two factors that have slowed the acceptance and application of computer modelling tools for addressing such sustainability challenges. First, the simulated agents in most social simulations are perceived as overly simplistic and inadequately realistic in their *psychological* architectures and network interactive rules. This leads to a reluctance among many social scientists and policy professionals to trust the outputs of such simulations. Such hesitation is understandable, since the agents and interactions in many models are indeed unrealistic, especially in the case of evolutionary game theoretic and economic models that assume humans are "rational actors." Even these simplistic models can in fact shed light on the micro- and meso-level causal mechanisms that lead to the emergence of macro-level phenomena, which helps to explain why (despite the reluctance in some corners) CMSS is rapidly gaining in popularity in both social science and policy analysis and evaluation [10–12]. Nevertheless, we argue that overcoming the reluctance mentioned above can be facilitated by implementing more realistic cognitive architectures and interpersonal networks within an approach sometimes called multi-agent artificial intelligence (MAAI) modelling [13]. In this approach, psychologically realistic architectures that link beliefs and behaviors can be combined with (without relying upon) machine learning techniques in order to simulate the emergence of population level trends from individual level (inter)actions.

The second factor is perhaps even more problematic. Even if we can identify the plausible mechanisms and feasible pathways that would enable us to navigate the Anthropocene, we do not yet seem capable of the sort of *collective action* that would be required to move us along said pathways. Applying intelligence (whether human or artificial) to these challenges is not enough; we also need to learn how to contest some of the evolved biases that render human decision-making and collective action across groups and cultures so difficult. In other words, we need to be mindful of those aspects of human nature that can be friction points to the adoption or implementation of well-informed results from complex social simulations. Among these biases is the tendency toward engaging in prosocial behaviors primarily or (under some conditions) only in relation to one's own in-group, whether kith and kin or a larger "imagined community" based on shared beliefs and norms.

What would it take to produce the conditions—and utilize the mechanisms—for collective action of the sort that would be needed to respond adequately to the challenges of the Anthropocene, which have contributed to what the 2024 UN Human Development Report framed as a "new uncertainty complex" characterized by a convergence of factors such as intensified polarization within and across countries, dangerous anthropogenic environmental changes, and sweeping societal transformations [14]? Whatever else it might take, making mid- and long-term progress toward the UN Sustainable Development Goals will require us to learn how to promote modes of prosocial individual and collective behavior that involve a healthier balance between merely *parochial* altruism (within groups) and a more *universal* altruism (across groups) within a larger proportion of the population.

This may sound improbable. It could turn out to be impossible. How might we find out? We recommend an approach that builds on recent advances in MAAI modeling and emphasizes working with a wide range of stakeholders, subject-matter experts, and change agents to model and simulate the conditions under which—and the mechanisms by which—altruism (of the sort that could foster widespread and sustainable cooperation, psychological and social well-being, and collective action) can be increased in empirically validated "artificial societies" that represent the relevant socio-ecological systems. The use of artificial societies can facilitate a more inclusive dialogue about the assumptions and aspirations that guide the (conceptual or computational) modeling of pathways through the Anthropocene. It can do this by facilitating the development and deployment of social simulation experiments that can test multiple theoretical hypotheses and practical policies before implementing them in the real world.

We use the phrase "sustainable altruism" to indicate an equilibrium at (and between) the individual and group level in which self-oriented concern for psychological and social wellbeing, on the one hand, and other-oriented prosocial attitudes and behaviors, on the other hand, can be achieved, maintained, and adapted as complex socio-techno-ecological systems continue to change during the Anthropocene [15,16]. The link between altruism and sustainability has often been observed and discussed in the literature [17–27], but our contribution here is intended to be novel both materially and methodologically. We highlight the complexity within and overlap between micro-, meso- and macro-level dynamics shaping prosocial attitudes and behaviors, describe the computational tools that are able to render that complexity more tractable (and thus explainable to policy experts), and offer recommendations for adequately attending to the ethical issues that emerge as we tackle the task of analyzing and facilitating pathways toward a more balanced and sustainable altruism within individuals and groups.

The three main sections of this article identify some of the evolved cognitive and coalitional biases that make it so difficult to achieve an equilibrium of sustainable altruism in contemporary human societies (CHALLENGES TO ACHIEVING SUSTAINABLE ALTRUISM (Section 2)), describe some of the innovative ways in which recent CMSS, and especially MAAI, approaches to psychological and cultural modeling open up new opportunities for simulating solutions to these challenges (OPPORTUNITIES FOR MODELING PROSOCIALITY AND PEACE (Section 3)), and address some of the ethical issues associated with using modeling and simulating methods to help us proactively navigate the Anthropocene (TOWARD A SIMULATION ETHICS FOR NAVIGATING THE ANTHROPOCENE (Section 4)). We conclude by summarizing the overall argument of the article, reiterating the challenges and opportunities related to participatory modeling approaches, and calling for complementary and collaborative efforts toward understanding and facilitating sustainable altruism.

CHALLENGES TO ACHIEVING SUSTAINABLE ALTRUISM

For reasons described in this first main section, it may well be that such an equilibrium does not exist in the possibility space of the human evolutionary landscape. For reasons described in the following main section, collaborative CMSS techniques appear to be the most promising set of exploratory tools we currently have for helping us find out whether there is an empirically plausible pathway to more widespread sustainable altruism. For reasons described in the final main section, even if these tools do not help us find such a pathway, we still have much to learn from an ethically attentive computational exploration of the possibility space of our shared future. Before jumping into these computational opportunities, however, it is important to be clear about the breadth and depth of the evolved cognitive and coalitional biases that render prosociality and peacebuilding so difficult. Only by surfacing these all-too-human tendencies and capacities can we hope to construct psychologically realistic models and pragmatic simulations for promoting the sort of balanced altruism that could enable collective action in response to current and coming challenges.

The Evolution of Human Prosociality

Natural selection favors organisms who act in ways that care for themselves long enough to procreate and pass on their genes (egoism). Eusocial species such as our early hominin ancestors, however, evolved in groups and in such cases natural selection also favored populations in which individuals sometimes cared for others in their group, even to the point of self-sacrifice (altruism). It makes intuitive sense that parents (or other genetically related individuals) would behave altruistically toward their children or close relatives because doing so increases the probability that (at least some of) their own genes will replicate in future generations. This is commonly referred to as *kinship* altruism [28,29]. We are not surprised when we see a caregiver exhaust his or her own energy and resources for the sake of an offspring.

However, we may be more surprised when we observe people behaving altruistically toward other non-kin individuals and even strangers, risking their own survival (and perhaps their own children's survival) by sacrificing time, food, or even their own lives to help. Many evolutionary biologists argue that this phenomenon can be explained by theories about trading networks, reputational management, or costly signaling that hypothesize that such behaviors indirectly increase the longer-term chances of one's own (or one's group's) genes surviving into the distant future. This is commonly referred to as *reciprocal* altruism [30,31]. Recent theories of cultural cybernetics suggest that this form of altruism can be expanded to include large-scale, imagined groups based on shared beliefs, motivations, or identity [32,33]. The phylogenetic inheritance (or moral equipment) that has been passed on to contemporary humans includes a capacity (and even a tendency under some conditions) to demote or even overcome conscious egoistic concerns and behave altruistically toward conspecifics.

In the social psychological literature, these dynamics are more commonly treated under the heading of prosociality, which typically refers more broadly to behaviors that are intended to benefit others whether or not such behaviors are intrinsically other-oriented or driven by underlying self-oriented concerns such as fear of punishment or hope for reward [34]. Most contemporary psychological theories and therapies are attentive to the way in which the evolutionarily stabilized tension between self and other structures attempts to find a balance between survival-based concerns for the self and healthy interpersonal connections with others [35]. A recent literature review outlined evidence for the growing consensus in psychology that concern for self and concern for others are intrinsically related developmentally and should be integrated into programmatic intervention strategies. Along with emotional regulation, empathy for self *and* empathy for others underpin the human capacity for prosociality [36].

At this stage, however, it is important to acknowledge the dark side of human altruism; namely, the tendency toward prosocial behaviors in relation to in-group members seems to have evolved alongside (and to be reciprocally related to) antagonism and aggression toward out-group members. In other words, the individuals who survived early ancestral environments (and passed on their genes to us) lived in social groups whose cohesion depended in part on their capacity to keep down the percentage of cheaters and freeloaders in their ranks by punishing those who betrayed the norms of the in-group or refused to participate in violence against out-groups when resources were scarce [37–41]. As we will see below, computer modelling has played an important role in testing hypotheses about in-group and out-group altruism. All of this goes a long way in explaining how human prosociality evolved in the late Pleistocene (approximately 130,000 to 11,500 years ago) as well as why the tension between individuals and groups in our contemporary environments makes social cohesion so fragile and peacebuilding so difficult in the early Anthropocene.

Parochial and Universal Altruism

Especially under stressful conditions, the positive tendency to care for the members of one's own group flips over to its own dark side, the negative tendency to feel anxiety toward or even attack outgroup members. This in-group oriented behavior, which is prosocial inwardly but antisocial outwardly, is commonly called *parochial* altruism [42–46]. While parochial altruism was clearly adaptive for our ancestors (otherwise they would not have survived) who lived in small-scale, hunter-gatherer societies, this trait has become increasingly maladaptive—or at least destructive—in our contemporary environment where most of us live in large-scale, pluralistic, population-dense, globally interconnected societies. In this new environment, which we as a species have helped to create (the Anthropocene), extreme expressions of parochial altruism make it difficult to live together with out-group members, which one encounters far more regularly, and contributes to conflicts and far more destructive wars, particularly when hoarding resources seems to be the only way for one's own group to survive.

The collective action required for surviving the Anthropocene seems to call for more *universal* altruism, the capacity to act prosocially in relation to the human species as a whole [47]. Or perhaps even extending altruism beyond our species to all sentience, life, the earth, or being itself [48–50]. The tendency or capacity to expand the range of altruism beyond one's own group varies in relation to individual psychological differences such as higher global consciousness, interoception, and non-religiosity [51–54], and appears to be promoted in secular cultural contexts that have higher existential security and broader social welfare programs [55–58].

The problem is that sustaining widespread universal altruism in any human population (much less the global population) appears to be highly unlikely and probably impossible. We know this from quantitative neuroscientific and biological research on the inherently "discriminating" and reciprocal structure of the evolved prosocial cognitive biases widely distributed in human populations and from qualitative data derived from psychological experiments, interviews, surveys, and ethnographic research [59,60]. We also know this from quantitative data from game theoretical computational experiments, which show that just a few selfish (non-cooperative) agents in a population consistently “win” over evolutionary time against altruistic agents [61–63], a topic to which we will return in the second main section of this article where we begin to tackle the issue of modeling and facilitating collective action more directly.

But we also know this from experience and common sense. We humans thrive in relatively small coalitions of family and friendship networks. It is hard to imagine a world in which parents do not discriminate at all in caring for their own children. Most of us might not even want to live in such a world. Some mystics and saints may appear to be able to achieve and maintain universal altruistic attitudes and behaviors, but they typically depend on the care of local “lay” resources and a wider network of parochially altruistic social systems. Parochial altruism is deeply embedded in our phylogenetic inheritance and has been socially entrained over millennia through communal rituals and hierarchical war-making organizational structures that simultaneously increase willingness to punish defectors and to kill out-group members in order to protect the resources and reputation of the in-group. These tendencies are activated and intensified under conditions of scarcity or threat that increase human anxiety about their own survival and the survival of those they care about [64–66].

Sustainable Altruism for the Anthropocene

Watching the daily news gives us no reason to suspect that the natural, predatory, contagion, financial, and social threats of the sort that activate and intensify human anxiety are likely to disappear anytime soon. They are only likely to increase as we struggle to respond to the “new uncertainty complex” alluded to above, which is generated in part by challenges related to ideological polarization within and across countries, anthropogenic environmental changes, and sweeping societal transformations across the globe. Given the dark side of parochial altruism, sticking with the status quo does not seem a viable strategy for promoting widespread individual and collective well-being in these uncertain times. Given the phylogenetic depth and ontogenetic reliability of the emergence of the cognitive and coalitional biases that drive care for self and kin, the prevalence of communal, institutional, and governmental structures that reinforce ingroup-oriented prosociality, and the demonstrated susceptibility of a population of cooperators to even a small number of defectors, an idealized world of universal altruists seems equally unsustainable.

Can we find and follow a pathway that would lead to an equilibrium of “sustainable altruism” suitable for surviving and perhaps even thriving as we navigate the new uncertainty complex that characterizes our current global environment? As noted above, we are using this phrase to refer to a stable but flexible balance between self-oriented concern for wellbeing (at the individual and group levels), on the one hand, and other-oriented prosocial attitudes and behaviors, on the other hand. Can our species achieve, maintain, and adapt a suitable prosocial equilibrium as socio-ecological systems continue to change during the Anthropocene? Such an equilibrium would have to work at both the individual and group levels as well as between them.

At the psychological level, this would mean (at the least) that individuals are provided with the resources and guidance to learn how to balance the self-care that is necessary for survival with the care for others that is necessary for pleasurable and fruitful human attachment. At the sociological level, this would mean (at the least) that groups are organized in ways that provide the internal bonds necessary for maintaining the identity of the group without triggering overly defensive or offensive strategies that turn violent. Finding an equilibrium between these levels would mean (at the least) that the needs of individuals for attachment and authenticity were somehow balanced with the needs of diverse groups to manage their identity and productivity in relation to other groups. All of this would likely require an extraordinary expenditure of resources and the emergence of a novel civilizational form analogous to the shift from hunter-gatherer to sedentary-agricultural or from archaic to axial age modes of material and cultural organization.

That all sounds very complicated. And not very likely to work. It may well be that pessimism gets the last word. But let's engage in a thought experiment. What would it take to find and, if it exists, to follow a pathway of sustainable altruism that balances the evolved needs of individuals and groups and the intricate interactions between them? We would need some way of analyzing and predicting the behavior of complex adaptive psycho-social-ecological systems that accounted for mechanisms at the micro-level, meso-level, and macro-level as well as the causal interactions among those levels. Most of us cannot hold all of these mechanisms and conceptual levels in our head all at once nor mentally predict dynamic changes within such massively complex and non-linear systems over time and under varying conditions. But maybe we don't have to. Techniques in computational modeling and social simulation have now advanced to a stage that may warrant a little optimism in our search for a pathway to sustainable altruism.

OPPORTUNITIES FOR MODELING PROSOCIALITY AND PEACE

In the following three sub-sections, we outline some of these advancements and describe ways in which such tools are already being used to understand and facilitate prosociality and peacebuilding. First, we discuss the potential as well as the limitations of traditional agent-based modeling (ABM) approaches to simulating altruism. Second, we provide some examples of the simulation of altruistic attitudes and behaviors using multi-agent artificial intelligence (MAAI) modeling, which is a subset of ABM that is differentiated by its focus on more psychologically realistic, socially networked agents and its emphasis on interdisciplinary theoretical integration. Third, we emphasize the importance of participatory approaches to modeling that engage diverse subject matter experts and stakeholders when attempting to construct policy-relevant models that can simulate pathways to social cohesion and peace and facilitate collective action. Throughout this section, we allude to some of the lessons our research teams have learned as they have tackled these core social challenges, which leads us into the discussion in the final main section about ethical concerns surrounding the use of these technologies in the simulation of, by, and for altruistic individuals and groups.

Evolutionary Game Theoretic Agent-Based Modeling

Because ABMs are explicitly designed to analyze and predict dynamics within complex adaptive systems, they are a particularly promising mode of social simulation. Unlike system-dynamics models, which simulate the dynamic causal flows between and among some variables (such as energy, information, or structures) within a social system, ABMs are also able to shed light on the changes within and among heterogeneous individual agents as they interact with one another and their environment, thereby linking micro-, meso-, and macro-levels of analysis in a way that no other social scientific method can. The potential implications for promoting prosociality within a human population are clear: well-constructed, empirically validated ABMs can illuminate the *conditions* under which—and *mechanisms* by which—variables such as social cohesion or peace at the macro-level can emerge from shifts in attitudes and behaviors among individuals and institutions at the micro- and meso-levels. We will return to this pragmatic potential in more detail below when we discuss policy-relevant participatory modeling strategies.

However, the scientific potential for understanding the evolution of human altruism and predicting changes in human prosocial behaviors is also clear. For this reason, it is not surprising that many early uses of ABM in social science addressed issues that are at least indirectly related to altruism [67,68]. For example, Schelling's model of segregation, perhaps the most well-known early social simulation, showed how easily widespread segregation can occur within a population when simulated individuals have even small preferences for living near in-group members [69]. However, it is widely agreed in the field that in order to take full advantage of the opportunity for ABM approaches to positively influence policy and shared collective action in response to shared challenges such as those related to sustainability and social cohesion, it will be necessary to develop far more realistic psychological agents and social network interaction architectures [70–78].

The most well-known and influential approaches to the use of simulation to understand altruism have been variations of evolutionary game theoretic modeling. Limiting ourselves to some recent examples, we can point to the use of this sort of ABM for modeling population-wide adoption of prosocial common-pool behavior [79], the role of an agent's own and their friends preferences in the formation of coalitions [80], the impact of influential individuals in promoting prosocial practices in heterogenous societies [81], and the extent to which diverse motives work together to motivate altruistic behaviors [82]. Unfortunately, game theoretic models are particularly susceptible to the charge of lack of psychological realism. They are evolutionary in the sense that agents pass on their (non)altruistic tendencies to new generations of agents, and they are game theoretic in the sense that this transfer occurs after agents "win" or "lose" during an interaction on a given run of a simulation. Such models are based on utility functions in which an agent calculates the potential benefit and cost of cooperating or defecting with others, i.e., engaging in altruistic or egoistic behaviors, which presupposes "rational actor" theory.

However, cognitive and social psychologists have known for decades that human behavior is not driven by rational calculations of utility, especially under stressful conditions that threaten survival, but by "bounded" rationality characterized by a variety of cognitive, emotional, and social biases [83,84]. Elinor Ostrom was one of the early critics of the limitations of game theoretic models and received a Nobel prize for her efforts to develop new models that more adequately accounted for the complex and networked dynamics of human behavior, which operate within nested clusters of human groups of various sizes [85–88]. While computational models and pragmatic proposals that have been inspired by Ostrom's work on the dynamics involved in the emergence of collective action have led to important insights [89–91], their focus on group selection and lack of adequate attention to the underlying role of evolved cognitive and coalition mechanisms leaves them open to the critique that they are psychologically unrealistic, which makes it difficult for scientists in other disciplines or public stakeholders to take them seriously.

Multi-Agent Artificial Intelligence Modeling

MAAI modeling attempts to remedy this problem by paying careful attention to the evolved psychological mechanisms that play a role in driving conflict or cooperation between human individuals or groups and to the value of incorporating multiple disciplinary perspectives into the cognitive architectures and interaction rules into computational models of social phenomena. For further discussion of this approach in general and of the rationale behind these two aspects of MAAI modeling strategy in particular see [13,15,92,93]. In this sub-section, we limit ourselves to a few examples that illustrate these aspects and the way in which they can contribute to increasing trust in—and therefore the usefulness of—computational models among other scientists and policy professionals. In our experience, when the causal architectures of social simulations implement multiple theories of human emotion and behavior that shed light on a phenomenon [94] and place the agents in realistic social networks informed by empirically validated evolutionary theories such as “Dunbar numbers” [95], other scholars and practitioners are more likely to find them plausible.

The dynamics of *parochial* altruism have been explored in a MAAI model that simulated the conditions under which—and the mechanisms by which—mutually escalating religious violence (MERV) emerges within a population [96]. The cognitive architectures and behavioral tendencies of simulated agents in this model were explicitly informed by social psychological theories such as terror management theory, social identity theory, and identity fusion theory, all of which have been empirically demonstrated to predict antagonism and violence (against outgroup members) using other methodologies. Each agent within MERV also belonged to a group and had variables that represented the evolved cognitive and coalitional biases that contribute to religious parochialism. Using real-world data from The Troubles in Northern Ireland (which lasted over three decades) and the Gujarat riots (which lasted over three days), MERV’s causal architecture was validated by simulation experiments that generated the emergence of the macro-level escalation of violence between two groups in the artificial society from the micro-level behaviors and interactions of its simulated agents (with over 98% accuracy).

Other MAAI models have more explicitly focused on *universal* altruism and broader prosocial attitudes and actions. For example, building on the Artificial Society Analytics Platform [97], which provides a virtual laboratory populated by agents with personality and other psychological variables in family, work, and online social networks, one model simulated the effects of religious belief and affiliation on prosociality [98]. Validated using longitudinal data from the World Values Survey, this model shed light on the roles that trust, tolerance, and religious group (non)affiliation played in promoting more universal (and decreasing parochial) altruism. Another model adapted the Artificial Society Analytics Platform to simulate the effect of religious pluralism (diversity of group beliefs) and religious credibility enhancing displays in the environment (prevalence of participation in in-group religious rituals) on disaffiliation in secular contexts [99]. A more historically oriented model implemented evolutionary and psychological theories about the function of anxiety and religiously motivated behaviors in promoting the kind of “prosocial equilibrium” required for the emergence and sustainability of large-scale societies [100].

It is important to note that all of this is relevant not only for addressing the scientific challenge of improving psychological and network realism, but also for addressing the pragmatic challenge of understanding and fostering the conditions and mechanisms that can promote prosociality and peace in contemporary contexts. Recent advances in machine learning and large language models have generated not only fear about AI but also some hope that AI might contribute to social cohesion and peacebuilding in ways we humans have not been able to imagine. However, human intelligence, learning, and use of language does not occur in a vacuum but is embodied and embedded in complex networked relations and emerges in response to affordances in the natural and social environment. This is why we believe that *multi-agent* AI holds perhaps even more promise for finding and following pathways toward a sustainable altruism that promotes collective action of a sort that can help us navigate the early Anthropocene.

Simulating Pathways to Peace Through Participatory Maai Modeling

In this sub-section, we briefly describe some recent MAAI models developed by our research teams that have attempted to move beyond simulating prosociality and altruism generally and toward simulating the conditions and mechanisms that could promote more sustainable peace within and among diverse human groups. Our goal has been to construct and explore the possibility spaces of validated models of some of the most complex and conflicted real-world socio-ecological systems. Over the years we have learned the importance of involving engaged stakeholders and change agents in the development, calibration, and validation of such models so that the outcomes are more likely to be relevant and useful for the concrete challenges they regularly face in the contexts in which they work [101,102]. Building on the growing number and quality of approaches to *participatory* computational modeling [92,103–105], we have been honing our own approach to focus particularly on the needs of governments and civil society organizations committed to addressing challenges related to sustainability in general and peacebuilding in particular.

For example, in the context of a 2022–2023 project with the Woolf Institute in Cambridge called "Forgiveness and Future-Building" we constructed a MAAI model of contemporary Northern Ireland in dialogue with a variety of stakeholders including change agents who had been involved in the original Good Friday Agreement. Members of the research team spent time talking with former combatants on both sides as well as experts and practitioners working on the ground to maintain and enhance the social cohesion and relatively peaceful situation in that post-conflict region. The context and some of the findings of this model are provided in a Background Paper for the 2023/2024 UN Human Development Report titled "Simulating Sustainable Societies: Uncertainty, Complexity and Multi-Agent Artificial Intelligence Modeling" [106], and a more detailed report is available in an article that was part of a special issue on the changing character of war and peacemaking [107]. Stakeholders in that context have been able to draw actionable insights from the results of the simulation experiments, which indicate (among other things) that the top two features that predict conflict are the number of values being discussed in the society and the differences in beliefs about those values between the two groups. These findings suggest that change agents in that environment could contribute by promoting discussion of a larger number of values in society, thereby providing more alternatives to the "sacred values" of the groups which, when challenged, can all to easily lead to violence among those individuals "fused" to those values [32,108].

During the fall of 2023, members of the same research team worked with the UNDP Program for Assistance to the Palestinian People to develop a prototype model that could provide policy-relevant insights for promoting the resilience of the Palestinian people. The development of this model, which led to what we called the "Palestine-Israel Virtual Outlook Tool" (PIVOT), was shaped by the input and ongoing feedback of a core group of UNDP staff and representatives from Palestinian and Israeli civil society organizations. Here too we spent time on the ground, visiting individuals and communities affected by the ongoing conflict. More advanced digital twins that were informed by this work are currently being used by other stakeholders engaged in attempts to resolve that conflict. In 2024 we worked with the UNDP in Bosnia and Herzegovina (BiH) to develop a model designed to shed light on the causes and consequences of horizontal trust (between diverse groups in the Balkans) and vertical trust (between the population in general and various institutions). The construction of the BiH model also involved a core group of experts and was guided by a larger advisory board that consisted of representatives from other UN agencies, civil society organizations, and academics located in Sarajevo.

The details of these last two models cannot be provided for security and privacy reasons, but an article describing the general approach of the PIVOT project [109] and a UNDP webinar describing some of the outcomes of the BiH project [110] can be found online. Our work in these and other regions continues, and we plan to report further on these processes in future publications. The main reason for discussing them briefly here is to highlight the importance and value of robust engagement with stakeholders and change agents before, during, and after the process of constructing MAAI models or any other computational social simulations. Only in this way can we hope that they will find the models trustworthy and utilize them in their peacebuilding or peacekeeping efforts. But this raises a host of questions. Who is invited to the (computer modeling) table? Which voices are excluded from that table? How will the simulation tools be used and by whom? If the goal is collective action, the process must include collective deliberation about the ethical ambiguities, assumptions, and aspirations at work in MAAI modeling.

TOWARD A SIMULATION ETHICS FOR NAVIGATING THE ANTHROPOCENE

In this final main section, we briefly outline some of the ethical implications and issues surrounding the development of policy-oriented social simulation in general, but especially MAAI models designed to promote social cohesion and peace in a way that might help us respond to the challenges of the early Anthropocene. We do not pretend to know all the complexities of the ethical landscape that will emerge nor how to navigate them in advance. Nevertheless, we have some recommendations for consideration as the social simulation field moves ever more deeply into the waters of policy-relevant peacebuilding oriented computational modeling. As Gilbert and colleagues have pointed out, whenever faced with difficult challenges that affect human well-being, it could be considered unethical *not* to use modeling methodologies that are explicitly designed to render complex adaptive systems more tractable [10].

We begin by highlighting the value of incorporating the findings of evolutionary biology and moral psychology about the ambiguity of human parochial altruism into the cognitive architectures and social networks of simulated agents. Next, we stress the importance of accounting for the fact that simulations designed and executed by human modeling teams will inevitably be susceptible to these same biases toward egoism and parochialism, and of finding ways to contest them. Finally, we note that this also applies to those stakeholders and change agents who utilize such models for some specific policy-oriented purpose. In other words, we call for surfacing and tackling all of these ethical issues head on in the process of constructing simulations of, by, and for altruistic individuals and groups.

Our first recommendation is to incorporate the *ambiguities* of human altruism in the construction of simulated artificial agents and the rules that guide their interactions with each other and their environment. Any attempt to develop realistic simulations that can illuminate the conditions under which—and the mechanisms by which—human individuals and groups might achieve an equilibrium of sustainable altruism should include the findings of the evolutionary and psychological sciences about the prevalence and potentially problematic nature of parochial altruism [111]. Given our inherited moral equipment, it is similarly important to acknowledge the implausibility of calls for widespread universal altruism. It makes sense to encourage people to broaden their sense of community or “we-ness,” but we should also acknowledge the human need for close relationships that provide a sense of identity. The challenge is finding the balance between caring for oneself and those in one’s own ingroup, on the one hand, and behaving prosocially within and beyond one's own ingroup, on the other. Participatory MAAI modeling can only help if it takes the ambiguities of altruism seriously.

But that won't be enough. Our second recommendation is to attend to the *assumptions* of those evolved humans involved in the modeling process itself. There may be no way to completely escape subjectivity biases and attain pure objectivity, but one of the desiderata of the scientific method is to open ourselves up to useful constructive criticism and correction so that we can become more aware of our own presuppositions and the surreptitious role of motivated reasoning in our own thinking. One of the values of CMSS in general and MAAI in particular is that the process of formalizing one's ideas, theories, and hypotheses about causal relations within a computational architecture is that it encourages and enables modelers to surface their assumptions. Among these assumptions are ideas about human nature and behavior, as well as attitudes toward the modeler's own ingroup(s) and whatever outgroup(s) they might experience as a threat to their own identity or well-being. Participatory modeling strategies can facilitate surfacing these assumptions and articulating them clearly, as well as providing a way of testing their coherence in relation to a broader artificial society and exploring their implications through simulation experiments [103,109,110].

Our third recommendation is to surface the *aspirations* of those who will be utilizing and applying the outputs of computational models and simulation experiments designed to inform the design and implementation of policies that will affect individuals and groups in the real world. In other words, it is important to pay attention not only to the assumptions built into a model but also the intentions of those who plan to use it. All of the stakeholders and change agents we have worked with have embraced ideals such as human rights, freedom, diversity, social cohesion, and peace. They would agree with us that one of the main goals of the modeling processes in which we have engaged is to promote what we have been calling sustainable altruism. However, like nuclear power, genetic engineering, or any other powerful technology, MAAI models could be used by "bad actors" for their own nefarious purposes. Such tools could also be used to discover the conditions under which—and the mechanisms by which—conflict could be increased in a population. Even peace-loving stakeholders may well disagree about the desired outcome that should guide the design and use of a model. There is no easy way around this. All models are purposeful abstractions, and the purposes of abstracting depend on the human modelers involved. The more intentionally we can surface and discuss differences in modeling aspirations among diverse stakeholders, the more likely we are to develop and deploy simulations that account for the social ethical dimensions of the challenges we face together [15,112–114].

CONCLUSIONS, IMPLICATIONS, AND FUTURE RESEARCH DIRECTIONS

In this article, we have argued for the potential value of participatory MAAI modeling for addressing the pragmatic challenge of promoting a sustainable altruism among individuals and groups that will enable us to find pathways toward more peaceful coexistence. Along the way, we have noted that this approach to modeling also addresses the scientific challenge within computer science and social simulation of creating more psychologically realistic artificial agents whose interactions occur in more realistic social networks.

We have also noted the limitations and challenges related to this and all CMSS approaches. As the saying goes, all models are wrong—but some are useful. They are all wrong because (like any map) they are purposeful *abstractions*, and abstractions do not include every detail. However, some are useful because (like a good map) they include those elements and structures that are most likely to successfully point the way to the desired outcome. As we have seen, another limitation has to do with their *purposefulness*, which raises a host of ethical questions including those discussed briefly above. Despite these challenges, we believe it is worth taking advantage of the opportunities made possible by advances in MAAI modeling to collaborate in exploring the possibility space of our species' future to see whether we can find a path to a prosocial and peaceful equilibrium (or equilibria) that can help us navigate these early years of the Anthropocene. What are some of the main takeaways for stakeholders and change agents interested in promoting sustainable altruism through participatory MAAI modeling or other CMSS approaches?

- Modeling human attitudinal and behavioral change in "artificial societies" requires simulated agents that are sufficiently psychologically realistic and interact with each other and their environment in realistic social networks. Capturing the complex, biased, and emotion-driven cognition of real humans in a computational architecture is hard work, but it is worth the effort because it enhances the plausibility of the models as well as their analytic and predictive power.
- Social simulations embed the assumptions and aspirations of their creators and users, which can lead to biased or ethically problematic outcomes if not made transparent. It is important to include diverse voices in the modeling process, which can help to mitigate against the use of models in ways that perpetuate imbalances of power and injustice. By involving a variety of stakeholders directly, participatory modeling helps surface minority perspectives, increase legitimacy, and foster buy-in for complex interventions.
- MAAI models enable researchers and policymakers to analyze nonlinear, multi-level interactions in complex socio-ecological

systems—something that traditional analytic methods struggle to manage. This helps to illuminate the way in which micro-level individual psychological traits and behaviors can scale up to macro-level societal outcomes such as cohesion, conflict, or sustainable cooperation.

- Policy makers can use such models as "virtual laboratories" for testing theoretical assumptions and policy peacebuilding interventions before implementing them in the real world. Simulation experiments can help change agents discover the conditions under which—and the mechanisms by which—mutual trust and peaceful coexistence can emerge, thereby offering strategic foresight for conflict resolution.
- Addressing contemporary local and global sustainability challenges will require the promotion of sustainable altruism—the capacity of individuals and groups to find a balance between behaviors that enhance their own survival (and thriving) and behaviors that promote forms of collective action for the common good. This will be incredibly difficult because all of us have inherited deeply rooted, evolved biases that foster parochial altruism. Participatory MAAI modeling may be one of our best tools for finding pathways to peace in the possibility space of the human species as we navigate our complex, intertwined, and rapidly changing societal and ecological environments in the early Anthropocene.

By leveraging psychologically realistic cognitive architectures and empirically grounded interaction networks, MAAI models provide a powerful lens through which to explore the dynamics of prosociality and the possibility space for peacebuilding. These tools enable both scholars and stakeholders to simulate and analyze how evolved cognitive biases and institutional configurations can either hinder or facilitate collective action, making them increasingly vital for addressing the global uncertainty complex we currently face as a species.

Where do we go from here? Several avenues for future research present themselves. First, more work is needed to integrate affective and cultural variables into agent architectures, particularly those related to anxiety, identity, and intergenerational memory, which play crucial roles in both prosocial development and conflict persistence. Second, efforts should focus on scaling and diversifying the participatory modeling process to ensure broader inclusion of marginalized voices and contexts, especially from the Global South. Doing so will enhance the representational validity of models and democratize access to scenario planning tools. Third, researchers must develop frameworks for ethical reflexivity within simulation design and deployment, ensuring that these technologies are used to empower rather than manipulate. Finally, planning and executing longitudinal studies that compare simulation predictions with real-world policy outcomes could significantly enhance the credibility and utility of MAAI for anticipatory guidance and social foresight.

As climate disruption, social polarization, and geopolitical instability continue to intersect, the need for rigorous, inclusive, and ethically grounded modeling of human cooperation grows ever more urgent. Despite its limitations, MAAI modeling offers a potentially generative pathway toward understanding and enacting sustainable altruism in our rapidly evolving global society.

DATA AVAILABILITY

No data were generated from the study.

AUTHOR CONTRIBUTIONS

Conceptualization, FLS; Writing—Original Draft Preparation, FLS; Writing—Review & Editing, FLS, JEL, MG, KO.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. McNeill JR, Engelke P. The Great Acceleration: An Environmental History of the Anthropocene since 1945. Cambridge (UK): Belknap Press of Harvard University Press; 2016.
2. Steffen W, Rockström J, Richardson K, Lenton TM, Folke C, Liverman D, et al. Trajectories of the Earth System in the Anthropocene. Proc Natl Acad Sci U S A. 2018;115(33):8252-9.
3. Cork S, Peterson GD, Janssen MA, Anderies JM, Elmqvist T, DeClerck F, et al. Exploring Alternative Futures in the Anthropocene. Annu Rev Environ Resour. 2023;48(1):25-54.
4. Ellis EC. The Anthropocene condition: evolving through social–ecological transformations. Philos Trans R Soc B Biol Sci. 2024;379(1893):20220255.
5. Lindahl T, Schill C, Lindahl J, Moore JC, Schlüter M, Biggs R, et al. Titanic lessons for Spaceship Earth to account for human behavior in institutional design. Npj Clim Action. 2024;3(1):56.
6. Beckage B, Lacasse K, Winter JM, Metson GS, Orians CM, Whitney CK, et al. Linking models of human behaviour and climate alters projected climate change. Nat Clim Change. 2018;8(1):79.
7. BenDor TK, Scheffran J. Agent-Based Modeling of Environmental Conflict and Cooperation. Boca Raton (US): CRC Press; 2018.
8. Bury TM, Bauch CT, Anand M. Charting pathways to climate change mitigation in a coupled socio-climate model. PLoS Comput Biol. 2019;15(6):e1007000.
9. Shults FL, Wildman WJ, Toft MD, Danielson A. Artificial Societies in the Anthropocene: Challenges and Opportunities for Modeling Climate, Conflict,

and Cooperation. Proc Winter Simul Conf. 2021;1-12. doi: 10.1109/WSC52266.2021.9715391
10. Gilbert N, Ahrweiler P, Barbrook-Johnson P, Narasimhan KP, Wilkinson H. Computational Modelling of Public Policy. J Artif Soc Soc Simul. 2018;21(1):1-19.
11. Hofman JM, Sharma A, Watts DJ, Stewart BM, Mislove A, Airoldi EM. Integrating explanation and prediction in computational social science. Nature. 2021;595(7866):181-8.
12. Lazer DM, Pentland A, Watts DJ, Aral S, Athey S, Contractor N, et al. Computational social science: Obstacles and opportunities. Science. 2020;369(6507):1060-2.
13. Lane J. Method, theory, and multi-agent artificial intelligence: Creating computer models of complex social interactions. J Cogn Sci Relig. 2013;1(2):161-80.
14. United Nations Development Programme (UNDP). 2023/2024 UN Human Development Report - Breaking the gridlock: Reimaging cooperation in a polarized world. 2024. Available from: https://hdr.undp.org/system/files/documents/global-report-document/hdr2023-24reporten.pdf. Accessed on 01 May 2025.
15. Shults FL. Simulation, Science, and Stakeholders: Challenges and Opportunities for Modelling Solutions to Societal Problems. Complexity. 2023;2023:1375004.
16. Shults FL. Gilles Deleuze and the Atheist Machine: The Achievement of Philosophy. Edinburgh (UK): Edinburgh University Press; 2024.
17. Borghesi S. Intergenerational altruism and sustainable development. 1999. Available from: https://cadmus.eui.eu/bitstream/handle/1814/725/1999_EUI%20WP_ECO_040.pdf. Accessed on 01 May 2025.
18. Halali E, Perez O. Binding the Future: Far-looking altruism boosts long-term sustainability. Available from: https://scholar.archive.org/work/pc7eiloypffsfhc6j34cylrw7m/access/wayback/https://files.osf.io/v1/resources/9szky/providers/osfstorage/65840165034461161468f4f3. Accessed on 01 May 2025.
19. Bréchet T, Lambrecht S. Altruism and sustainability: a framework for analysis. 2004. Available from: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fc19b1ccf0f9bf61f9bcfaaaeca1c93373b44213. Accessed on 01 May 2025.
20. De Dominicis S, Schultz PW, Bonaiuto M. Protecting the environment for self-interested reasons: Altruism is not the only pathway to sustainability. Front Psychol. 2017;8:1065.
21. Fossas-Tenas A, Ibelings BW, Kasparian J, Krishnakumar J, Laurent-Lucchetti J. Paradoxical effects of altruism on efforts to mitigate climate change. Sci Rep. 2022;12(1):13072.
22. Schäpke N, Rauschmayer F. Going beyond efficiency: including altruistic motives in behavioral models for sustainability transitions to address sufficiency. Sustain Sci Pract Policy. 2014;10(1):29-44.

23. de Morais LHL, Pinto DC, Cruz-Jesus F. Circular economy engagement: Altruism, status, and cultural orientation as drivers for sustainable consumption. Sustain Prod Consum. 2021;27:523-33.
24. Romeiro AR. Sustainable development and institutional change: the role of altruistic behavior. Texto Para Discussão. 2000. Available from: http://www.eco.unicamp.br/images/arquivos/artigos/1727/texto97.pdf. Accessed on 01 May 2025.
25. Schill C, Schlüter M, Lindahl T, Moore JC, Orach K, Wijermans N, et al. A more dynamic understanding of human behaviour for the Anthropocene. Nat Sustain. 2019;2(12):1075-82.
26. Tapia-Fonllem C, Corral-Verdugo V, Fraijo-Sing B, Durón-Ramos MF. Assessing sustainable behavior and its correlates: A measure of pro-ecological, frugal, altruistic and equitable actions. Sustainability. 2013;5(2):711-23.
27. Tripathi RC, Singh S. Psychosocial Pathways Towards a Sustainable Society: The Role of Greed, Altruism and Social Capital. Psychol Dev Soc. 2017;29(2):200-20.
28. Dawkins R. The Selfish Gene: 30th Anniversary Edition. 30th ed. Oxford (UK): Oxford University Press; 2006.
29. Harman O. A history of the altruism–morality debate in biology. Behaviour. 2014;151(2–3):147-65.
30. Trivers RL. The evolution of reciprocal altruism. Q Rev Biol. 1971;46(1):35-57.
31. Kurzban R, Burton-Chellew MN, West SA. The Evolution of Altruism in Humans. Annu Rev Psychol. 2015;66(1):575-99.
32. Lane JE, Shults FL, Wildman WJ. A potential explanation for self-radicalisation. Behav Brain Sci. 2018;41.
33. Lane J. Cultural Cybernetics of Religion: Computation and Information Transmission in Religion. In: The Routledge Handbook of Evolutionary Approaches to Religion. London (UK): Routledge; 2022. p. 409-25.
34. Malti T, Speidel R. Prosocial cascades: Understanding and nurturing the potential for positive developmental trajectories. Adv Child Dev Behav. 2023;64:189-216.
35. Homme GA, Shults FL. The Shaping of the Self: Patterns and Pathways in Bowlby, Kohut, and Bowen. Aust N Z J Fam Ther. 2020;41(4):367-80.
36. Malti T, Speidel R. Development of prosociality and the effects of adversity. Nat Rev Psychol. 2024;1-12.
37. Purzycki BG, Apicella C, Atkinson QD, Cohen E, McNamara RA, Willard AK, et al. Moralistic gods, supernatural punishment and the expansion of human sociality. Nature. 2016;530:327-30.
38. Bourrat P, Atkinson QD, Dunbar RIM. Supernatural punishment and individual social compliance across cultures. Relig Brain Behav. 2011;1(2):119-34.
39. Fehr E, Gächter S. Altruistic punishment in humans. Nature. 2002;415(6868):137-40.

40. Krasnow MM, Delton AW, Cosmides L, Tooby J. Group Cooperation without Group Selection: Modest Punishment Can Recruit Much Cooperation. PLoS One. 2015;10(4):e0124561.
41. Rusch H. The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension. Proc R Soc B Biol Sci. 2014;281(1794):20141539.
42. Abbink K, Brandts J, Herrmann B, Orzen H. Parochial altruism in inter-group conflicts. Econ Lett. 2012;117(1):45-8.
43. Choi JK, Bowles S. The coevolution of parochial altruism and war. Science. 2007;318(5850):636-40.
44. Sheikh H, Atran S, Ginges J, Wilson L, Obeid N, Davis R. The Devoted Actor as Parochial Altruist: Sectarian Morality, Identity Fusion, and Support for Costly Sacrifices. Cliodynamics. 2014;5(1):23-40.
45. Bernhard H, Fischbacher U, Fehr E. Parochial altruism in humans. Nature. 2006;442(7105):912-5.
46. Kustov A. Borders of Compassion: Immigration Preferences and Parochial Altruism. Comp Polit Stud. 2021;54(3–4):445-81.
47. Epstein RA. Altruism: Universal and Selective. Soc Serv Rev. 1993;67(3):388-405.
48. Zwick M, Fletcher JA. Levels of Altruism. Biol Theory. 2014;9(1):100-7.
49. Ma HK. The development of altruism with special reference to human relationships: A 10-stage theory. Front Public Health. 2017;5:271.
50. Österblom H, Paasche Ø. Earth altruism. One Earth. 2021;4(10):1386-97.
51. Piech RM, Strelchuk D, Knights J, Hjälmheden JV, Olofsson JK, Aspell JE. People with higher interoceptive sensitivity are more altruistic, but improving interoception does not increase altruism. Sci Rep. 2017;7(1):15652.
52. Liu JH, Choi SY, Lee IC, Leung AK, Lee M, Lin MH, et al. Behavioral evidence for global consciousness transcending national parochialism. Sci Rep. 2023;13(1):21413.
53. Tsang J-A, Al-Kire RL, Ratchford JL. Prosociality and religion. Curr Opin Psychol. 2021;40:67–72.
54. Galen L, Sharp M, McNulty A. Nonreligious Group Factors Versus Religious Belief in the Prediction of Prosociality. Soc Indic Res. 2015;122(2):411–432.
55. Zuckerman P. Society without God: What the Least Religious Nations Can Tell Us About Contentment. New York (US): NYU Press; 2010.
56. Shults FL. Practicing Safe Sects: Religious Reproduction in Scientific and Philosophical Perspective. Leiden: Brill Academic; 2018.
57. Paul G. The Chronic Dependence of Popular Religiosity upon Dysfunctional Psychosociological Conditions. Evol Psychol. 2009;7(3):398–441.
58. Norris P, Inglehart R. Sacred and Secular: Religion and Politics Worldwide, 2nd edition. Cambridge (UK): Cambridge University Press; 2011.
59. Hardin G. DISCRIMINATING ALTRUISMS. Zygon®. 1982;17(2):163–186.
60. Reinhardt L, Whitehouse H. Why care for humanity? R Soc Open Sci. 2024;11(4):231632.

61. Braver SL, Barnett B. Effects of modeling on cooperation in a Prisoner's Dilemma game. J Pers Soc Psychol. 1976;33(2):161.
62. Alexander JM. Evolutionary game theory. 2023. Available: https://www.cambridge.org/core. Accessed on 02 Jan 2025.
63. Ibrahim AM. The conditional defector strategies can violate the most crucial supporting mechanisms of cooperation. Sci Rep. 2022;12(1):15157.
64. Greenberg J, et al. Evidence for Terror Management Theory II: The Effects of Mortality Salience on Reactions to Those Who Threaten or Bolster the Cultural Worldview. J Pers Soc Psychol. 1990;58(2):308–318.
65. Mcgregor HA, et al. Terror Management and Aggression: Evidence That Mortality Salience Motivates Aggression Against Worldview-Threatening Others. J Pers Soc Psychol. 1998;74(3):590–605.
66. Bonin K, Lane JE. Identity fusion, devoted actor theory, and extremism. In: Routledge Handbook of Evolution and Religion. New York (US): Routledge; 2022.
67. Hodder I. Simulation studies in archaeology. In: New directions in archaeology. Cambridge (UK): Cambridge University Press; 1978.
68. Doran J, Palmer M, Gilbert N, Mellars P. The EOS project: Modelling Upper Palaeolithic social change. In: Gilbert N, Doran J, eds. Simulating Societies: The Computer Simulation of Social Phenomena. London (UK): UCL Press; 1994:195–221.
69. Schelling TC. Dynamic models of segregation. J Math Sociol. 1971;1(2):143–186.
70. McNamara JM. Towards a richer evolutionary game theory. J R Soc Interface. 2013;10(88):20130544.
71. Jager W. Enhancing the realism of simulation (EROS): On implementing and developing psychological theory in social simulation. J Artif Soc Soc Simul. 2017;20(3). doi: 10.18564/jasss.3522
72. Hofstede GJ, Frantz C, Hoey J, Scholz G, Schroder T. Artificial Sociality Manifesto. Rev Artif Soc Soc Simul. 2021. Available: https://rofasss.org/2021/04/08/artsocmanif/. Accessed on 05 Jan 2025.
73. Tzafestas E. Modeling Meta-Cultures. Available: http://www.softlab.ece.ntua.gr/~brensham/Publications/ESSA2011.pdf. Accessed on 05 Jan 2025.
74. Gotts NM, et al. Agent-based modelling of socio-ecological systems: Models, projects and ontologies. Ecol Complex. 2019;40:100728.
75. Schaat S, Jager W, Dickert S. Psychologically plausible models in agent-based simulations of sustainable behavior. In: Agent-Based Modeling of Sustainable Behaviors. Cham (Switzerland): Springer; 2017; pp. 1–25.
76. Mitsopoulos K, et al. Psychologically-valid generative agents: A novel approach to agent-based modeling in social sciences. In: Proceedings of the AAAI Symposium Series. 2023:340–348. Available: https://ojs.aaai.org/index.php/AAAI-SS/article/view/27698. Accessed on 05 Jan 2025.
77. Wijermans N, Scholz G, Neumann M, Paolillo R, Templeton A. Social identity modelling. JASSS. 2023;26(3):15.

78. Costantini A, Scalco A, Sartori R, Tur EM, Ceschi A. Theories for computing prosocial behavior. Nonlinear Dyn Psychol Life Sci. 2019;23:297–313.
79. Sotnik G, Shannon T, Wakeland W. A new agent-based model offers insight into population-wide adoption of prosocial common-pool behavior. J Math Sociol. 2023;47(4):311–338.
80. Kerkmann AM, Cramer S, Rothe J. Altruism in coalition formation games. Ann Math Artif Intell. 2024;92(3):601–629.
81. Crabtree SA, Wren CD, Dixit A, Levin SA. Influential individuals can promote prosocial practices in heterogeneous societies: a mathematical and agent-based model. PNAS Nexus. 2024;3(7):224.
82. Wu X, Ren X, Liu C, Zhang H. The motive cocktail in altruistic behaviors. Nat Comput Sci. 2024;4(9):659-676.
83. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science. 1974;185(4157):1124–1131.
84. Kahneman D. Thinking, Fast and Slow, Reprint edition. New York (US): Farrar, Straus and Giroux; 2013.
85. Ostrom E, Burger J, Field CB, Norgaard RB, Policansky D. Revisiting the Commons: Local Lessons, Global Challenges. Science. 1999;284(5412):278–282.
86. Ostrom E. A general framework for analyzing sustainability of social-ecological systems. Science. 2009;325(5939):419–422.
87. Ostrom E. Collective action and the evolution of social norms. J Econ Perspect. 2000;14(3):137–158.
88. Ostrom E. Do institutions for collective action evolve? J Bioeconomics. 2014;16(1):3–30.
89. Kliemt H. Tayloring Game Theory the Ostrom Way. Good Soc. 2011;20(1):37–49.
90. Frischmann BM, Marciano A, Ramello GB. Retrospectives: Tragedy of the commons after 50 years. J Econ Perspect. 2019;33(4):211–228.
91. Wilson DS, Philip MM, MacDonald IF, Atkins PW, Kniffin KM. Core design principles for nurturing organization-level selection. Sci Rep. 2020;10(1):13989.
92. Diallo S, Wildman WJ, Shults FL, Tolk A, eds. Human Simulation: Perspectives, Insights, and Applications. Berlin (Germany): Springer; 2019.
93. Shults FL. Progress in simulating human geography: Assemblage theory and the practice of multi-agent artificial intelligence modeling. Prog Hum Geogr. 2022;46(1):108-120.
94. Shults FL, Lane JE, Diallo S, Lynch C, Wildman WJ, Gore R. Modeling Terror Management Theory: Computer Simulations of the Impact of Mortality Salience on Religiosity. Relig Brain Behav. 2018;8(1):77–100.
95. Gore R, Lemos C, Shults FL, Wildman WJ. Forecasting changes in religiosity and existential security with an agent-based model. J Artif Soc Soc Simul. 2018;21:1–31.
96. Shults FL, Gore R, Wildman WJ, Lynch C, Lane JE, Toft M. A Generative Model of the Mutual Escalation of Anxiety Between Religious Groups. J Artif Soc Soc Simul. 2018;21(4):1-25.

97. Shults FL, Wildman WJ, Diallo S, Puga-Gonzalez I, Voas D. The Artificial Society Analytics Platform. In: Advances in Social Simulation. Cham (Switzerland): Springer; 2020.
98. Galen L, Gore R, Shults FL. Modeling the Effects of Religious Belief and Affiliation on Prosociality. Secul Nonreligion. 2021;10(1).
99. Cragun R, McCaffree K, Puga-Gonzalez I, Wildman W, Shults FL. Religious exiting and social networks: computer simulations of religious/secular pluralism. Secul Nonreligion. 2021;10(1).
100. Puga-Gonzalez I, Shults FL, Gore R, Talmont-Kaminski K. An agent-based model of prosocial equilibrium: Simulating the role of religiously motivated behavior in the formation and maintenance of large-scale societies. In: Advances in Social Simulation. Cham (Switzerland): Springer; 2023.
101. Lane JE, Shults FL. Cognition, Culture, and Social Simulation. J Cogn Cult. 2018;18:451–461.
102. Shults FL. Simulating theory and society: How multi-agent artificial intelligence modeling contributes to renewal and critique in social theory. Theor Soc. Available from: https://link.springer.com/10.1007/s11186-025-09606-6. Accessed on 14 Feb 2025.
103. Ahrweiler P, Gilbert N, Pyka A. Joining complexity science and social simulation for innovation policy: agent-based modelling using the SKIN platform. Cambridge (UK): Cambridge Scholars Publishing; 2016.
104. Mehryar S, Sliuzas R, Sharifi A, Reckien D, van Maarseveen M. A structured participatory method to support policy option analysis in a social-ecological system. J Environ Manage. 2017;197:360–372.
105. Shults FL, Wildman WJ. Human Simulation and Sustainability: Ontological, Epistemological, and Ethical Reflections. Sustainability. 2020;12(23):10039.
106. Shults FL, Lane JE, O'Lone K. Simulating Sustainable Societies: Uncertainty, Complexity and Multi-Agent Artificial Intelligence Modelling. UNDP 2024 Human Development Report Background Paper. 2024. Available online: https://hdr.undp.org/system/files/documents/background-paper-document/2023flshultsjelanesimulatingsustainablesocieties.pdf. Accessed on 01 May 2025.
107. O'Lone K, Gantley M, Lane JE, Shults FL. Employing Multi-Agent AI to Model Conflict and Cooperation in Northern Ireland. N Engl J Public Policy. 2024. Available: https://scholarworks.umb.edu/nejpp/vol36/iss1/6/. Accessed on 06 Sep 2024.
108. Sheikh H, Ginges J, Coman A, Atran S. Religion, group threat and sacred values. Judgment Decis Making. 2012;7(2):110–118.
109. Gilbert D. The UN Hired an AI Company to Untangle the Israeli-Palestinian Crisis. WIRED; 2023. Available: https://www.wired.com/story/culturepulse-ai-israeli-palestinian-crisis/. Accessed on 26 March 2024.
110. UNDP Innovation. Global Innovation Series | Harnessing AI for Social Cohesion in Bosnia & Herzegovina. 2024 Aug 15. Available: https://www.youtube.com/watch?v=ogym2xJTwEc. Accessed on 06 Jan 2025.

111. Shults FL, Wildman WJ. Ethics, Computer Simulation, and the Future of Humanity. In: Diallo SY, Wildman WJ, Shults FL, Tolk A, eds. Human Simulation: Perspectives, Insights, and Applications. Cham (Switzerland): Springer International Publishing; 2019; p. 21–40.
112. Diallo SY, Shults FL, Wildman WJ. Minding morality: ethical artificial societies for public policy modeling. Ai Soc. 2021;36(1):49–57.
113. Shults FL, Wildman WJ, Dignum V. The ethics of computer modeling and simulation. In: 2018 Winter Simulation Conference (WSC). New York (US): IEEE; 2018; p. 4069–4083.
114. Shults FL, Wildman WJ. Artificial Social Ethics: Simulating Culture, Conflict, and Cooperation. Proc SpringSim 2020 Conf. 2020:1–10.

How to cite this article:
Shults FL, Lane JE, Gantley M, O'Lone K. Sustainable Altruism: Simulating the Future of Prosociality and Peacebuilding—A Conceptual Review. J Sustain Res. 2025;7(2):e250033. https://doi.org/10.20900/jsr20250033