16

What Is Collaboration Good For?

Raul Pacheco-Vega, Rebecca D. Hardin, A. C. C. Coolen, George E. Marcus, Chris Nierstrasz, Justin Sulik, Sander E. van der Leeuw, Paul F. M. J. Verschure, and Alicia von Schenk

Abstract

This chapter explores collaboration from the perspective of its purpose, driving factors, and outcome: why we collaborate, how we undertake it, and what goals are pursued through our joint efforts. It begins by analyzing factors that contributed to the evolutionary success of our species—pressures that shaped us to collaborate intuitively and inherently. It explores how collaboration is linked to the challenges of managing common pool resources and distinguishes between utilitarian from vicarious forms. It stresses the limitations of economic models, which overly emphasize rational self-interest, and argues to replace the reductionism prevalent in economic theories with a nonlinear approach. Approaching future research this way will enable a more nuanced, comprehensive understanding of collaboration—one that transcends simple transactional relationships and incorporates a deeper, more integrated view of human behavior and interaction.

What Is Collaboration?

Over hundreds of thousands of years, *Homo sapiens* have established themselves as quintessential social beings. The preeminent factor contributing to human evolutionary success lies in our species' flexible ability to adapt and shape diverse physical and sociocultural environments, an adaptability that has facilitated the conquest of the planet. This adaptability critically depends on the collective intelligence or "social mind," which emerges from the interplay of numerous individual minds through symbol systems and language. Over time, intragroup communication and intergenerational knowledge transmission have fostered the expansion and refinement of the social mind. Consequently, these

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

processes, which have bolstered human adaptability and promoted social development and progress, are rooted in prosocial abilities. The reciprocal interaction between adaptation and prosocial behavior, therefore, has shaped our trajectory as sophisticated social and collaborative organisms.

Collaboration is born out of inherent human impulses or instinctual behaviors; it takes a sustainable form when a challenge surpasses that which a single actor can meaningfully tackle. As such, it can extend beyond a specific group or type of actors, human or otherwise. Collaboration implies a complementarity among the participants, involving a recursive interplay of diverse skills among multiple, heterogeneous agents. While diversity is crucial, there is also an intermediary phase of alignment in which shared beliefs are sought and collective objectives are formulated, leading to mutual conceptualizations of the collaborative process. Indeed, heterogeneity is an indispensable prerequisite for collaboration, particularly in relation to the burgeoning forms of collaboration that involve human–human, human–nonhuman, and wholly nonhuman agents.

Collaboration provides both direct and indirect advantages to the interacting agents. Direct benefits are typically symbiotic, creating mutual benefit, whereas indirect gains encompass a more interconnected and, hence, stable local and global community. Much of the fruitfulness of these collaborative efforts emerges through the creation of specialized domains or niches, contributing to an overall metastability. This process of niche creation is an emergent outcome of recurrent collaborative interactions.

When addressing complex problems through collaborative means, strategies need to be devised that conceptualize, construct, and aspire toward establishing *islands of stability* amidst the intricate dynamics of our physical and social environments. At the same time, such strategies should not be too deeply embedded in those temporary islands of stability because the very solutions we produce can change the environments we inhabit, begetting new problems to address. One intrinsic characteristic of successful collaborations is their ability to balance stability and adaptability. How is this equilibrium embedded within the objectives and architectures of collaborative processes, and is that maintenance of metastability a primary purpose of collaboration?

A significant harmonizing element in collaboration rests in the mental capability to project future objectives onto present contexts at varying levels of abstraction, subject to different premises or possibilities. In the context of a collaborative group, the target being sought is merely an imagined result, a potential narrative reflecting a prospective future, or, in other words, a virtualization (see Chapter 5, this volume). These virtualizations are anchored in communication, reinforced by collective action, and institutionalized within societal norms or systems. They are built from primitive and ancient cognitive processes, such as the capacity for mental time travel (conceptualizing past or future scenarios) and mentalizing (empathizing with and understanding the minds and intentions of others). Collaboration by necessity occurs in both the real world and an imagined one.

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

In the context of collaborative efforts, an envisioned and thus virtualized outcome might take the form of a tangible depiction of the problem at hand, the desired outcome of its resolution, or the strategy the group plans to employ to tackle the issue. An envisioned outcome, however, might rapidly become obsolete as the environment evolves, due to inherent changes or the group's actions. This may result in a higher discrepancy when comparing the predicted future state to its actuality. Such an eventuality, or prediction error, could potentially create risk and the dissolution of the collaboration or precipitate a loss of trust in each agent's capabilities to attain shared objectives. Conversely, an envisioned outcome might sometimes be more abstract and act more as a unifying force rather than a tactical guide (e.g., "Liberty, Equality, Fraternity!" or "Brexit means Brexit!"). This unifying force capitalizes on human sociality, particularly the human predisposition that collaboration is constitutive of our cognitive and social lives. If the state of the world shifts, an abstract virtualization remains useful as a rallying cry, even if it is does not suggest a map of future trajectories which could be called "progress."

By virtualizing goals at different levels, both as motivations for action and as organizing forces, flexibility as well as stability are possible in collaboration. One could contend that specific mechanisms for these exist (e.g., humor and ritual) in collaborative processes. A second balancing force is the cognitive diversity within a group.

In a dynamic and complex world, problem-solving and collaboration can involve a vast number of dimensions. However, human cognition is limited; it is able to attend to and process a small set of dimensions at any time. Different collaborations will thus involve different forms of dimensionality reduction by participating agents and their organization. Forms of collective and individual dimension reduction will also depend on the traits of the individual agents that make up the collaborating collective. If individuals are cognitively homogeneous, a problem may be reduced to similar dimensions across group members. There will be more resonance across their perspectives: they will be better able to communicate and to exploit more efficiently their current position in the problem space. However, if they are cognitively diverse (with different information-processing and behavioral styles, attentional patterns, or values), then the dimensionality reduction will be correspondingly more diverse across group members. Their perspectives will be more disjointed: they will be better able to conceive of new strategies and better positioned to adapt if the state of the world shifts dramatically. Hence, collaborative processes balance exploration and exploitation by converging on different configurations of collaboration defined through the predispositions of the agents involved, the information architecture of the collaborative process, and the properties of the niche or task.

Collectively, the quality of the forces that drive collaboration indicates that collaboration can operate on multiple temporal scales across various metastable states. Diverse virtualizations and dimensionality reductions provide unique

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

methods for individuals and groups to influence one another and to comprehend the task at hand as well as each other. Social norms can modulate agents' freedom to manage the nonlinear states of their interactions with one another and the world. For example, norms might allow for the adaptive strengthening or weakening of social bonds or for a spike in creativity and innovation during crisis periods. These forces further offer assorted perspectives on the nature of collaboration itself, driven by overarching, abstract virtualizations and shared viewpoints of the problem. Yet, collaboration can also become an end in itself, as in play. Team members establish and validate their membership through interactions, thereby building networks of trust. However, when a crisis or challenge arises, the emphasis shifts from collaboration as an end in itself to collaboration as a means to achieve mutually aligned objectives. We note that self-referential collaboration is where specialization and a division of labor can be created. Such forms of interaction can have positive transformative effects on the efficiency and power of collaborative processes but also raise questions about how control over the collaboration and its utility is created, shared, or lost. We thus argue that collaboration is both realized and exercised also for the sake of collaboration itself, in the absence of an external need, problem, or goal. Indeed, we can distinguish *utilitarian* from *vicarious collaboration*. The former is shaped through external reinforcement, whereas the latter results from an intrinsic motivation to collaborate. Utilitarian collaboration focuses on what is; vicarious collaboration is shaped by what could be by virtue of the agents' ability for virtualization (see Chapter 5, this volume).

Any collaboration is predicated on a sense of shared futures or destiny, Margaret Levi explains (Ahlquist and Levi 2013; Chapter 6, this volume). This requires implicit optimism that such a future can exist and be obtained through the collaborative process, bringing positive outcomes to the parties involved. Collaboration thus requires the ability of mental time travel where multiple scenarios of possible futures can be imagined, valued, and communicated.

As much as agents are biased by their priors, systems of collaboration also have preconditions, including institutions that help shape collaborative processes through norms and conventions. Trust, morality, and intentionality are all part of this set of preconditions. This implies that collaboration can be seen as a multiscale system comprising multiple substrates, including agents, tasks, institutions, norms, and environment. The collaborative process occurs through multiple feedback loops within and between these scales, forming an architecture of collaboration (see also Chapters 1, 5, and 14, this volume).

What is the objective of collaboration? One way to articulate the collective aim of collaboration is to conceptualize it as the collective pursuit of creating pockets of stability while maneuvering through a complex world. This clarifies why we strive to collaborate. It is important to acknowledge, however, that due to continuous feedback loops, shifting group membership and structure, and the characteristics of open, dynamic systems, no two collaborative experiences

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

can ever be identical. To paraphrase Heraclitus, *we can never step into the same collaboration twice.*

If we accept that the benefits of human collaboration (and hence the reasons for human agents to collaborate) are not always purely practical but sometimes strictly social (i.e., that collaboration also serves to satisfy an innate desire of humans to interact with other humans), then the benefits of these interactions will inevitably be mutual. A need felt by humans is being satisfied by the act of collaboration, which suggests that the cognitive apparatus must be able to track the quality of collaborations. Even when collaboration appears to be driven by altruism, one could still argue that there is a mutual benefit. The recipient of altruism experiences direct practical benefit, whereas the altruist tends to derive a sense of worth or satisfaction from the altruistic act, or the satisfaction of having served a higher moral cause. To quote Abraham Lincoln, "When I do good, I feel good; when I do bad, I feel bad. And that is my religion." It is hard to imagine a situation where an agent enters a collaboration voluntarily from which no benefit is derived, be it a practical outcome, social pleasure, or the satisfaction of doing something meaningful. The notion of inclusive fitness provides a biological underpinning for the tendency to collaborate for the sake of collaboration itself, the process itself becoming the goal.

Humans are open systems for matter, energy, and information construction and exchange. They thus have a built-in capacity but need to exchange information with their environments. In that sense, they are no different from other life forms, from single cellular to more advanced life forms. The exchange of information is based on resonance between the internal milieu and the niches constructed by information processing and action. Trust results when there is a sufficient degree of resonance between an individual and another, or an individual and a group of individuals; more specifically, the resonance between the internal models that agents maintain of each other. Do the predictions of agents and tasks that a collaborative agent generates align with the signals it receives? Trust is thus built on virtualization and predictability. Trust-building allows the partners concerned to increase the volume of information they exchange. Once that volume has grown sufficiently and thus created a degree of affinity between the parties' cognitive structures, they can attain a state of collaboration, which is a (partial) alignment between their visions, expectations, goals, and values. Trust-building allows collaborating agents to build a tacit understanding. The essence of that trust is that all collaborating agents expect an understanding of each other, or empathy, and on that assumption, trust the other's interpretations and reactions. Trust provides an example where collaboration for the sake of collaboration facilitates the co-creation of cognitive structures that support future collaboration by building trust (see also Chapter 11, this volume).

Through collaborations, we construct transient stability in a permanently dynamic process governed by various feedback loops. In this, agents always bring something to the table, and their mental models potentially change the

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

rules and parameters of collaboration. As soon as a collaboration starts, it will, by necessity, change the mental models, rules, and norms that govern the collaboration, which simultaneously become prerequisites and consequences.

Utilitarian versus Vicarious Collaboration

The recursive nature of collaboration can be applied to the canonical notion of collaboration as linked with the challenges of managing common pool resources (CPRs), defined as non-excludable and rivalrous (Ostrom 2015). Ostrom's study of communities managing CPRs was an empirical rebuke to the economics and public policy literature that restricted governance options to markets or states. Her design principles suggest conditions under which communities can successfully manage a common pool resource in the absence of a third-party enforcer (a coercive state) or an invisible hand (a "free" market). Recent research on the governance of CPRs shows that it creates multiple, measurable outcomes in what are actually complex biosocial systems (Chhatre and Agrawal 2009). Analyzing them with adequate disaggregation and synthetic insight requires the harmonization of multiple empirical studies (Ferraro and Agrawal 2021).

In one view, the commons are material or immaterial domains where the rules and norms that have been established are distinct from those for managing the adjacent or surrounding domains (see Chapter 12, this volume). Generations of scholarship have shown that the functioning of human rules and norms can be highly flexible, and property regimes or systems can be ill-defined and overlap or interpenetrate. Still, at their core, common property regimes to manage shared resources are domains where private, state, or open-access norms are not fully applicable. They are the domains of collaborative monitoring, use, and enforcement surrounded by more restrictive or permissive rules and norms that pertain to human and societal engagement. During the Middle Ages in Europe, commons were areas (e.g., central grazing areas for animals in the center of a village) that were open for use by community members according to norms agreed upon and enforced by the community, in contrast to all other areas for which there were no rules (open access) or use was determined by individual or group ownership (private property) or restrictive use (state ownership). Each governance structure is, however, exclusive and discrete.

Following Ostrom's leadership in the field of political science, natural resource-related examples have come to be seen as a repeated process of co-creation by communities to limit encroachment and overconsumption of a common pool resource. They have contributed to more robust legal and political understandings of property regimes and to more nuanced economic understandings of different types of goods, as they group along a spectrum of excludability in use and rivalry in consumption of exhaustible resources (Figure 16.1).

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

	Excludable	Non-Excludable
Rival	Private goods *"Typical goods"* (clothes, food, flowers)	Common goods *"Common Pool Sources"* (mines, fisheries, forests)
Non-Rival	Club goods *"Artificially scarce Goods"* (cable TV, private parks, cinemas)	Public goods *"Collective Goods"* (air, news, sunshine)

Figure 16.1 A typology of goods, suggesting where they fall on a spectrum of excludability and rivalry in consumption or the possibility of resource exhaustion.

The nuanced ideas that have come to undergird our understanding of governance, property, trade, and resource management constitute crucial social and natural science frameworks for conceptualizing collaboration. In general terms, they place collaboration in the context of a means-end analysis. Too often, the intuition has become: "Wherever there is collaboration, there must be a commons." In centering collaboration on common pool resource governance, Ostrom's influential position can be described as Aristotelean in that it makes stability the basic phenomenon of interest to understand collaboration, resting on a homeostatic regulation through negative feedback (for further discussion, see Chapters 5 and 11, this volume). This reductive framework, however, hardly captures the elegance of her attention to the elaboration of rules and norms, surveillance and self-surveillance, monitoring and sanctioning.

What of a unifying, dynamic theory of collaboration? Analysis of the early stages of colonial trade in Asia shows how governance structures can be hybrid and morph over time along a coercive/collaborative continuum (Chapter 3, this volume). The same may hold for trends in technologically enhanced collaboration. Seen thusly, collaboration toward the maintenance of a common pool resource is only one specific form of human collaboration. It must be conceived in relation to a more generic case that merely rests on the coordination of action in the service of collaboration itself or even collaboration as the commons among which collaborations arise and are maintained.

Collaborations are inscribed in a wider stream of diverse ways of working to construe islands of stability within oceans of change and flux. Many forms of human collective action display an alignment of action in the absence of a clearly identifiable common resource. We can see this, for instance, in dance, music, rituals, and team sports as well as in human-coordinated complementary actions that pursue goals that seem elusive or abstract, such as convincing the electorate to vote or to participate in education. By construing goals as dynamic, flexible, and multiple realizable virtualizations, the commons (and its role for collaboration) can at times dissolve and may even reemerge in a changed form. This perspective coincides with an overall Heraclitean

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

dynamism based on allostasis and stability through continuous change while striving for better generalizability (see Chapters 5 and 12, this volume).

How Does Collaboration Happen?

Alignment implies the creation of shared horizontal, co-constructed understandings, meanings, and actions. A useful metaphor is the island of stability, which emerges from the alignment among the participating agents in otherwise dynamic contexts. Obviously, collaboration requires this alignment to be present as a prior or to occur through repeated interactions among agents. Such interactions can arise from the sheer impulses and preferences of human agents, as we are an inherently collaborative species for whom interaction toward shared goals can be deeply gratifying. However, we also need to remember that each agent has their own priors regarding traits, values, and beliefs. Alignment does not need to produce a perfect match or fit, but it does require shared understanding between agents of the collaboration's (possibly evolving) goal. This understanding is shaped and possibly stabilized by mapping individual and shared virtualizations of the collaborative process into a common frame of reference. As much as agents are defined by their priors, shared narratives communicated to participating agents may shape their individual virtualization of the task at hand and amplify alignment.

In the case of vicarious collaboration structured around virtualized and compressed models of the triad of world, self, and tasks (see Chapter 5, this volume), what mental capabilities must agents have for collaboration dynamics to arise and stabilize? Agents must be able to align their actions in terms of both their coordination in time and space and their complementarity. This requires that each agent maintain a model of the other agents in the collective as well as in the task at hand. This can be captured under the general notion of theory of mind or mentalizing (i.e., the ability to maintain models of the other in terms of their internal states such as goals, knowledge, emotions, and skills) (Flavell 1999). In other words, virtualization comprises both the task and social spaces of the participating agents.

Models of Collaboration

The above description shows that collaboration emerges from a confluence of factors. To disentangle these, researchers have engaged with a wide range of collaboration models (for an overview, see Chapters 5 and 15, this volume). A good model can account for the relevant features of the process it seeks to describe while leaving out irrelevant details. The art—some might say "dark art"—is then to identify these relevant features. One method relies on domain knowledge, common sense, the objectives of the modeling exercise, and the modeler's experience to determine the relevancy of features. In the context of

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

collaboration, if the number of agents is large, nonlinear processes describe the agents' states and if these agents are heterogeneous (i.e., they are not clones), then we would call such models complex systems. To predict future behavior involves tricky mathematical analysis or extensive computer simulations (or both). The complexity increases even more when the model in question describes connected processes that operate on distinct timescales. Since the 1970s, significant progress has been made to expand our arsenal of mathematical tools with which to study complex systems.

Next, we discuss a particular element that seems to us both a definitive feature and a value of collaboration: the dynamism of systems. Complex systems with many variables tend to exhibit certain generic features, which suggests to us what to expect from them. Whenever these systems have positive feedback (or reinforcement) loops, and the level of randomness in the interaction processes is not too large, such systems tend to exhibit multiple local stable states or stable trajectories. Yet, this generates nontrivial dynamics; states or trajectories that will evolve from the system depend on initial conditions. Moreover, over time, the noise in the system may trigger spontaneous transitions from one metastable state, or manifold, to alternative (more stable) states or trajectories. Collaborating agents provide a good example of such a complex system (see above discussion on islands of stability).

Remanence

A further consequence of positive feedback in complex systems is remanence. This term describes the phenomenon whereby if we change the system's parameters to such an extent that an initially stable state of the system becomes unstable, and then revert back to our original parameter values, this does not mean that our system will also return to its initial stable state. We may well find ourselves on a new plateau, from which we can only escape through more drastic action. This is exactly what is expected to happen with climate change: in all probability, the damage we are inflicting on Earth will not simply be undone by restoring greenhouse gasses to preindustrial levels. The same may hold for the dynamics of collaboration.

A second feature of complex systems is that whenever the number of interacting variables is sufficiently large, even very simple local rules of interaction may give rise to highly nontrivial cooperative macroscopic phenomenology that could never arise and cannot be understood in small systems. These are sometimes called emergent properties; they are not explicitly "coded" into the rules of engagement of the interacting elements.

Large interacting systems can also show profound transitions between qualitatively different macroscopic states if we change system parameters. Freezing or melting of liquids or solids at very specific reproducible temperatures (crucial for life) would never happen in small systems: fridge magnets would not exist, computer displays would not exist, Earth's magnetic poles would not be

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

(meta)stable, proteins would not fold, autoimmune diseases would not be triggered from one day to the next. Models in which the agents represent decision-making humans rather than molecules, atoms, heteropolymers, or immune cells are not different in this respect. Also here, simple rules of interaction can lead to complex and counterintuitive collective phenomena and profound transitions if system parameters cross critical values. The same mechanisms are at work to cause crashes in financial markets or violent revolutions in societies. Hence, emergent properties can be menacing and unmanageable.

Impact of Interaction on Agent Features

While the features of agents clearly have an impact on their interactions, interactions also have an impact on the characteristics of agents. This is immediately obvious when evolutionary dynamics are considered: agents with features that increase their fitness to survive in the real world (where their ability to collaborate with other agents will certainly play a role) are more likely to survive and reproduce, which shapes the features of subsequent generations. However, the effects of interaction can also act on shorter timescales, as when agents learn or adapt their behavior in response to interactions with others. For all practical purposes, the adapted agent will be a different agent than the novice, which we can refer to as niche construction. This raises issues of ontogeny and phylogeny whereby collaboration constitutes us as internally differentiated communities of shared adaptive practice, distinct from others. From this perspective collaboration can be cast in terms of niche construction, where the collaborating agents create an effective space of interaction serving a commonly created goal (see also Chapters 2 and 12, this volume).

Dos And Don'ts in the Modeling of Interacting Human Agents

What are good models to describe interacting humans? It is far easier to describe what should not be done than to describe what should be. One of the historical pitfalls (especially among economists) is disregarding reality and allowing model choices to be led too much by considerations of mathematical convenience and conceptual simplicity. If one assigns to all human agents a perfect, rational, deductive mind; perfectly quantifiable objectives; and the possession of complete and perfectly accurate information, then one can model decision making by these agents using standard Nash-like game theory (Nash 1950). In combination with further assumptions, such as the Efficient Market Hypothesis, this leads to models of financial markets that are described by relatively simple and clean mathematical laws (Malkiel 2003). These laws predict that markets are always efficient and stable, describing a fantasy world in which humans do not actually live; in that fantasy, there would never be unfair advantages, crooks, price manipulations, stock market crashes, or wars.

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

It is hard to see the benefit of such model studies beyond academic sustenance. A more sensible alternative strategy was proposed by the economist Brian Arthur (1994, the El Farol Bar problem), streamlined by Challet and Zhang (1997, the Minority Game), and extended by many others. In these models, agents behave more like real-life humans: imperfect actors seek individual profit in a simplified competitive market.

In the Minority Game, agents are assumed to make money when they are in the minority group; in a market where sellers are the majority, the buyers and sellers do well. In the El Farol Bar problem, the agents are not in possession of complete information; they act inductively (i.e., they "try" different personal action strategies, not all of which are sensible, and select the ones that at any point in time appear to work best for them). At times, the agents behave irrationally, suffer from greed and herding sentiments, and may even decide to quit if they get frustrated. They are much closer to the imperfect species that humans tend to be. The most salient features of these models are that they demonstrate how even selfish individual actions by agents can lead to aggregate behavior that looks like the result of clever collaboration. They also exhibit phenomenology (e.g., avalanches, crashes, and highly nontrivial fluctuations) that are much closer to what is observed in real markets than in the synthetic world of the Efficient Market Hypothesis. In addition, Minority game-type models can still be understood mathematically.

To understand human interaction and collaboration using mathematical and computational models, one needs to go further. The Minority Game illustrates what is possible, and minority models show that one can quite effectively model the action of real and imperfect humans. By so doing we can see the predictability and volatility of complex systems, such as collaborating humans, and thus pose better questions for further research. In sum, we pose four major observations:

1. Nonlinear complex systems can display metastable states and remanence.
2. Large systems of multiple interactive agents can have final emergent complex ontologies.
3. Macroscopic collaboration can arise from microscopic competing agents.
4. Heterogeneity in complex systems is not random.

Toward What Is Collaboration Trending?

Above, we took a broader view of collaboration beyond the governance of a commons and were not content with emerging behaviorist lenses for conceptualizing collaboration, with their emphasis on targeted external feedback. We emphasized that agents bring their priors (e.g., history, theory of mind, virtualization) to the table. Here, we move away from emphasizing an external

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

perspective (where the environment restricts and shapes interactions to impose stability) and consider agents as co-creating a task space that includes or constitutes the commons at its core. We move our thinking from theories of a society adapting to or appropriating and managing elements of the environment, to one that coevolves with and irrevocably changes (or even co-creates) the environment, and hence conditions of possibility for future collaborations. We posit that this can be generalized to collaboration and interactions with nonhuman species and synthetic intelligence.

Both historical and future scenarios offer cautionary signals with respect to the sustainability of collaboration. Collaboration unfolds at various scales, from the individual and the group to that of institutions and societies and the collaborative task itself. When considering the psychosocial dynamics of collaboration, at least two scales come into play. Each creates their own risks for collaboration, and require dedicated arbitration and conflict resolution mechanisms to overcome.

First, when common goals or shared intentionality drive the collaborative relationship, there is a perpetual process of cognitively matching the common goal with the task environment in which the group operates. At each point in time, the possible paths over which the group can reach a future state through collaboration are reevaluated depending on the group's current state. As a result, discrepancies can and will occur between previously virtualized objectives and those encountered. These discrepancies will create stress on the collaborative process and may lead to its collapse, a novel view on its goals and strategy, or an altered mapping of actions to outcomes.

Second, individuals within the group continuously match their own goals with the group's overall performance and goals to monitor how these correspond. In case of a mismatch, individuals might leave the group or, even worse, act out of self-interest against the original common goal, thus again affecting the overall dynamics of collaboration or threatening its integrity. In general, the alignment between collective and individual goals needs to be maintained.

The success of a collaborating collective rests on its capability to co-construct shared narratives, virtualizations, goals, and objectives to maintain belief in these constructs. If we assume that collaboration interaction dynamics can be realized around virtual mental constructs, which are highly abstract and reflect states of the real world and its agents, how are interaction dynamics maintained in the absence of direct reinforcement (e.g., as provided by a community with shared reliance on a common pool resource)? We consider that various types of contemporary commons can be constituted by something as ephemeral as software, hence created as novel phenomena, then progressively elaborated and enlarged through collaboration rather than merely being socially constructed, managed, and governed as for material resources.

In our view, trust is a crucial shaping force that can modulate collaboration dynamics from the moment it commences. One of the priors that collaborating agents bring to any collaboration are their internal models. Trust is established

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

and maintained when these are supported by the actual validity of the triad task, self, and others. The consistency between these models and the world, therefore, defines trust and rests on the implicit and explicit ability to monitor the alignment of internal models and the things they describe. If trust is the reinforcer that maintains or collapses the mental structures supporting collaboration and their ensuing actions, we are invited to consider its constituent factors beyond the direct matching between internal models and external things and events. One such factor is group membership, which hinges on the bonding and affirmation an agent receives from the collective. This implies that the creation of groups and group membership form part of the arsenal of vicarious collaborative acts that prepare collectives for further collaborative acts of increasing complexity and volatility, at wider spatiotemporal scales.

A second factor can be the normativity of the structures within which collaboration unfolds (e.g., legal, cultural, or religious frames), which might be able to compensate for individual tendencies to disengage due to a lack of trust. In this case, trust in the process itself is superseded into trust in a higher-order ontological frame of reference (e.g., the law, ancestors, or a deity). These potentially powerful processes of identification and motivation are not to be confused with adaptive niche formation. They could be considered potential negative niche formation, with deleterious consequences for the field of collaborative action (see also Chapter 4, this volume).

Our detailed discussions of the Dutch East India Company's engagement with trade in Asia, toward the end of the eighteenth century, illustrate how stable collaborations may not be just, socially symmetrical, nor even a relationship between equal and similar actors. These co-created outcomes offer a way to study collaboration under a wide range of historical, social, and structural circumstances. Although each may pursue different strategies of coordination and coercion, they all constitute sustained, significant collaborations (see Chapters 3, 4, and 6, this volume). New tools have emerged to make collaboration easier to dismantle, distort, and destabilize.

The fear that global collaboration could be deliberately disrupted (e.g., for reasons of food, water, and energy scarcity) and exacerbate social unrest seems reasonable. Yet so is guarded optimism. There may be no better wide example of how current and past media are woven together and received in variously subversive ways than the exuberant, interruptible practice of *rickrolling*. Designed to short-circuit attention-grabbing algorithmic tendencies (or the inclination to take anything online too seriously), rickrolling speaks to a playful, difficult-to-predict set of impulses. In other words, if technological change is accelerating and amplifying, so too are mechanisms for its disruption, and not only in ways that produce chaos but also in ways that produce communities and, indeed, playful, self-reflexive, critical impulses among humans, even in widely disparate locations. To rely on adolescent attention spans and communities of morally motivated hackers (increasingly subject to the co-opting efforts of states and companies) seems preposterous. Yet, we are approaching

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

such scenarios where hybrid-collaborating collectives may shape the future because established structures of global and local collaboration are losing their effectiveness. Going forward, such scenarios must lend urgency and creativity to our investigations of biological, hybrid, and synthetic collaboration.

What Is a Commons and Why Does It Matter?

In her work on the management of CPRs, defined as non-excludable and rivalrous (Figure 16.1), Ostrom challenged the traditional notion that CPRs (e.g., forests, fisheries, and irrigation systems) are inevitably subject to overuse and depletion due to the tragedy of the commons. She argued that successful CPR management can be achieved by developing institutions that enable users to cooperate and coordinate their actions (see Chapter 13, this volume). Her research drew on a wide range of case studies of different types of CPRs from around the world and has influenced diverse fields, from environmental policy to development economics.

Her study of communities that manage CPRs was an empirical rebuke to the economics and public policy literature, which had led to a restriction of governance options to markets or states (Ostrom 2015). Ostrom's governance design principles suggest the conditions under which communities can successfully manage a CPR in the absence of a third-party enforcer (a coercive state) or an invisible hand (a "free" market). History shows, however, that governance structures can be more complex and change over time along a coercive/collaborative continuum (Chapter 4, this volume).

The Commons, Property Rights, and Digital Rights

It is potentially useful to distinguish between the commons and property rights. The economists' understanding of the commons considered historical precedents as well as mathematical models of processes that describe interacting human agents (collaboratively or otherwise).

Historically and now, in the virtual realm, common or privately controlled resources can emerge from open-access practices. These are areas or spheres of activity for which no coherent rules yet exist about how to exploit, own, or restrict the use of resources; for instance, the domain of action of the HSBC Bank in Hong Kong or Shanghai before banking was constrained by local rules; open source software innovation before it goes behind a paywall or gets integrated into platforms that are open only to subscribers; the information sphere where large tech companies roam before rules governing the Internet and associated cognitive spaces existed and users established their own rules.

"Open" frontiers (both virtual and real)—where commons can be established or private and public rights can be asserted—have become a frictional, generative frontier zone. Here, an unknown, potential commons confronts

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

individuals and society at the limits of the zones they have made into resources, by imposing rules and norms of human action and interaction. The resulting dynamics reveal connections between the cognitive niche and un-cognized environment that leads to the identification of resources and the development of ways to utilize these by individuals or society. It is a substrate for self-referential aspects of collaboration, constantly changing due to proliferating, relational collaborative practice that produces new patterns, and through this, also new possibilities (Krause and North 2020).

We do not use the term *frontiers* lightly; it resonates with violence at various points in human history (e.g., colonialism, trade, and missionary work suffused with military force) and refers to spheres of activity where a specific dynamic is ongoing and the outcome is not given. A frontier indicates potential for change: violence and alienation are possible but so too are innovation and new forms of collaboration. Borders, on the other hand, signify a static delimitation, based on distinctions (whether natural or human), that separates entities and enables enforcement. Whereas frontiers are relational, borders are "demarcational." Borders define specific criteria from among the many that apply to the space concerned and impose those as definitive. Borders are thus virtualizations that confer stability. Frontiers are unstable conjunctures out of which forms of sustained collaboration can emerge.

These conceptual distinctions acknowledge that not all effective collaborations improve the world from a given normative perspective. A collaborative process can work with authenticity, accountability, and impact without necessarily making the world a better place. This is easily recognized in the example of the Dutch East India Company (see also Chapter 3, this volume). After creating a global commons by declaring the freedom of the seas (*Mare Liberum*), Dutch traders and East Asian local rulers affixed their collaboration through contracts that defined agreed-upon goals, outcomes, and enforcement mechanisms (see Chapter 3, this volume). In the absence of a single monarchical authority, this turned out to be essential as it provided a way for the two parties to resolve potential problems. If one party defaulted, the other party could punish the breach of contract and obtain compensation to repair the damages, even if this meant using physical force and violence. This new set of parameters guided global geopolitics and trade, and had the potential to make the world a more stable place. Indeed, during this period of history, the notions of sovereign and sovereignty became more complex (Hardin 2011). Viewing collaboration from a historical perspective demonstrates how collaborations that take place within a frontier at one moment in history can give rise to conventions, regulations, and institutions that create borders or boundaries in a subsequent era.

The Case for Trepidation

Artificial intelligence (AI) and algorithms shape our ideas of medical, emotional, and educational well-being, generating hope as well as misgivings

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

about unintended long-term consequences of collaboration. We cannot remove such fears from what Robert Paehlke (2003) refers to as "democracy's dilemma" of electronic capitalism. Biological, hybrid, and synthetic collaborations are being shaped by a rapidly widening human use and concomitantly narrowing human control over machines. This paradox poses a serious and growing risk in our ability to forge collaborations and thus form islands of stability in what may be increasingly volatile oceans of flux (environmental and political).

One can easily imagine worst-case scenarios as well as constructive measures that would strengthen our abilities to connect, collaborate, and confer stability. For instance, the distinction between commons defined by property rights and those lying outside of it (as in some virtual realms) opens up an intellectual space to understand that limiting property rights is potentially desirable (e.g., for collaborations that enable humans to confront common challenges). It is appropriate, for example, to require that users be given control over their own data as a necessary condition for potential collaboration mediated by algorithms. Data inferred about an individual by a third party (e.g., the algorithmic profiling of a person) should be governed by the same rules as those governing privileged professional relationships; that is, such data may be used only in the individual's interest. Also, data that requires authentication by third parties for transaction purposes (authenticated personal data) and its distribution should also be in the hands of the individual whom this data describes. Such authenticated personal data would ensure that everyone has a unique digital identity and establish trust. Building on this, data that individuals wish to share among specific groups for specific collaborative purposes should be put into "data commons," in which the managers of these commons have a fiduciary duty to use the data solely in the interests of the members of the commons. We see these as initial steps to address emerging regimes of property rights in the domain of digital collaboration, which is a critical frontier of our time.

Ulrich Beck (1992) has long cautioned that the interests of the scientific and technical elite, who manage the design, production, programming, parameterization, research, and publishing about machines and their intelligence, are worth scrutinizing. These interests are not necessarily determined by corporate backers and donors but are certainly in dialogue or in collaboration with them. Such backers and donors have long been valuable allies to those engaged in scientific inquiry that is not certain to produce a profit, but which intuition and inference tell us might have intrinsic value. This situation is neither static nor necessarily negative; it does, however, describe a context in which collaborations are shaped.

Collaboration among a narrow but powerful global elite could be one of the largest threats to collaboration's natural diversity and proliferative tendencies. Increasingly, violence and even wars emerge from the expansion of security states, which often camouflage their grasp for power in response to

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

tensions among affinity groups (e.g., Russia's claim that it is fighting Nazis in Ukraine or U.S. authoritarian tendencies to use a moral imperative of freedom to protect against putative violent immigrants). Confrontations are escalating in frequency and intensity as evidenced from (a) pro- and anti-government clashes in parts of Central and South America, Hong Kong, or Sri Lanka, (b) friction between left- and right-wing ideologies in parts of Europe and North America with increasingly religious overtones, and (c) religious conflicts in both Africa and India that pit religious groups against each other for purposes of regional territorial, economic, and social control. This primitive violence is taking hold in economies with high technological and educational capabilities and representative political processes—surely one of the most complex forms of persistent collaboration at scale. But let us not overlook the strategic production of violence, where forms of sustained, expanding collaboration threaten the interest of elites, and thus incentivize their collaboration to counteract it. For instance, the collusion of early capitalist scions with Nazis in Germany has been under-documented and gone largely unpunished, even though they enabled and profited from the violence (de Jong 2022). A similar dynamic can be observed in the emerging oligarchies of Eastern Europe and the former Soviet Union.

To be reflexive and intellectually honest, we must also consider cases where the forces balancing collaboration lead to periods of stability, and thus collaborations produce more abstract information-processing superstructures: institutionalized strategies for gathering, filtering, or processing information as it becomes relevant to the collaboration, its objectives, and virtualizations. A core example is modern science itself: It is not necessarily part of the world but rather a cultural product that evolved over time to affix one human way of making sense of our experience of the world, as a way to control it. Science is part of a growing island of stability that has become an anchor for further and more extended collaborations, entrenching its own stability through institutions and norms. When taken too far in the direction of stability, however, it can become scientism, with various perverse effects. For instance, it could fail to achieve its goals and thus doom the collaborative enterprise. It can also produce a range of abuse, from eugenics or crimes committed in Nazi Germany, to the unethical and harmful Tuskegee syphilis experiments. These compromise the capability of science to stand for collaborative authenticity, humility about complex phenomena, and accountability among participants.

Such negative outcomes can result from appropriating collaborative practice for harmful political ends. However, it can also result from the practices and goals being virtualized at an inappropriate level of abstraction, thereby tending toward vicarious collaboration that pursues a master narrative devoid of operational relevance relative to the original collaborative process. An extreme example of such vicarious collaboration would be the actions of a mob driven by conspiracy thinking, reacting to a vision of the future that is a mere illusion.

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

The Case for Cautious Optimism

In diverse arenas, novel constructive forms of human collaboration build on new technologies. In the digital age, social media platforms have become pivotal in orchestrating large-scale human collaborations, particularly in addressing social, environmental, and political challenges. Notable instances include the *#MeToo* movement, which used social media to amplify voices against sexual harassment globally, and the *Ice Bucket Challenge*, which raised unprecedented awareness and funds for amyotrophic lateral sclerosis through viral challenges. The *Fridays for Future* movement, initiated by Greta Thunberg, leveraged platforms like Instagram and Twitter to mobilize millions of students for global climate strikes, while the Sunrise movement significantly impacted the U.S. Green Deal by mobilizing American youths, thus demonstrating the power of digital tools in environmental advocacy. Similarly, the *Black Lives Matter* movement expanded its reach and impact by strategically using social media to organize widespread protests and maintain global attention on racial justice. The *Arab Spring* provides another critical yet tragic example, where protestors coordinated and disseminated their efforts through social media, leading to significant political change across North Africa and the Middle East. *March for Our Lives* showcased how young people used digital platforms to organize one of the largest youth-led protests in American history, advocating for gun control. These examples highlight the transformative role of social media in facilitating effective and expansive collaborations, making it a fundamental component in contemporary social movements and campaigns.

Advances in AI have also changed how humans collaborate with algorithms as well as with each other. The realm of Go, an ancient strategy game, experienced a groundbreaking transformation when DeepMind's AI, AlphaGo, defeated world-class human players in 2015, nearly five years ahead of what experts had predicted. This pivotal event not only challenged the traditional paradigms of the game but also inspired a reevaluation of strategies and teaching methods among the Go community. Some Go masters, disillusioned by the prowess of nonhuman intelligence, chose to step away from competing and teaching, while others embraced the new insights offered by AI and began to integrate these advanced strategic concepts into their own gameplay and pedagogy, pushing human players to explore far deeper into the game's possibilities. This evolution has led to a vibrant new phase in the game of Go, where human and nonhuman intelligence collaborate, leading players to reconsider unviable moves and expand the game's strategic depth. This interaction between human skill/intuition and AI's search depth illustrates a unique and flourishing form of collaboration and competition, marking a significant shift in how the game is played and taught. Similar developments can be observed in other game environments. Indeed, computer game environments have become critical benchmarks in advancing the capabilities of AI systems.

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

Conclusion

Fear that collaboration can be undermined or nefariously directed seems reasonable in our present-day world. Continued ascendance of monolithic big tech, amid shortages of food, water, and energy scarcity, in turn, exacerbate social unrest and limit patience with the messiness of the representative political process, even as it is under attack by authoritarian-leaning leaders. Yet guarded optimism is also reasonable—even critical. Societal risks posed by social media loomed large in our conversations as coauthors and collaborators, but also as parents, educators, and entrepreneurs. Harm seems especially acute for adolescents. However, many young men and women are using technologies to co-create virtual worlds or to manage effective social movements, joining forces with those of other generations to organize, educate, and attain political traction on issues that matter.

To understand what collaboration is good for, we need to move past traditional economic models of collaboration and assume a more holistic approach by integrating psychological, social, and evolutionary perspectives. Economic models are limited as they overly emphasize rational self-interest and focus on altruistic and intrinsic motivations that drive collaborative behavior. By considering the evolutionary roots of collaboration, we may discover the pressures that shaped us to collaborate. The reductionism prevalent in economic theories must be replaced by a new class of models and theories that acknowledge the complex, dynamic nature of real-world collaboration, where the nonlinear dynamics of collaborative efforts and interactions are not straightforward and can lead to unpredictable outcomes. Approaching future research this way will enable a more nuanced, comprehensive understanding of collaboration—one that transcends simple transactional relationships and incorporates a deeper, more integrated view of human behavior and interaction.

From "The Nature and Dynamics of Collaboration,"
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144

From “The Nature and Dynamics of Collaboration,”
edited by Paul F. M. J. Verschure et al. Strüngmann Forum Reports, vol. 33,
Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262548144