1

Introduction

How Collaboration Arises and Why It Fails

Paul F. M. J. Verschure

Background

Through its ability to collaborate, *Homo sapiens* has become the dominant species on this planet. Traces of this ability can be observed in the remains of Neolithic construction projects, such as Göbekli Tepe from 9000 BCE (Turkey) with its mysterious large-scale circular structures. Comparable collective building projects on each continent characterize this period. The earliest written records describe the ability of humans to overcome their differences and work together toward common goals, as in Homer's *Iliad* (800 BCE), where Greek leaders needed to put aside their differences to achieve the shared objective of conquering Troy and recovering Helen. Millennia of human collaboration have left an indelible mark on Earth systems, ushering in the Anthropocene. This era, characterized by global biogeochemical, technological, economic, and sociocultural alterations, has led to unforeseen yet undeniable stress on populations, ecosystems, and planetary systems. Both the COVID-19 pandemic and present global conditions—from armed conflicts to escalating humanitarian crises and political polarization—demonstrate amply why we need to understand the phenomenon of collaboration. This knowledge may well be critical to our long-term survival. The persistent success of human collaboration is now creating conditions for its failure, suggesting that we might have reached the end of the road for human collaboration as we know it and must reinvent it following our current predicament.

The stability and sustainability of social and Earth systems depends on realizing stable collaboration dynamics. Yet how, and by which means, collaboration is achieved and maintained is not clearly understood; neither are the conditions nor processes that lead to its breakdown or failure. Indeed, as this volume illustrates, the exact meaning, connotation, and use of "collaboration" are unclear, as is its difference from notions like "cooperation" or "collusion."

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 this distinction, differences between intent, morality, legality, and outcomes matter. This suggests that we do not yet have the proper concepts to describe and understand this powerful, complex, orchestrated form of human collective behavior. Ironically, as the need for effective collaboration is more urgent than ever, we lack a common understanding of the concept. This is a common problem at the frontiers of knowledge and not automatically an invitation to a potentially endless concept analysis. Still, it shows that we have barely scratched the surface of understanding this mission-critical phenomenon, highlighting the need for collective reflection and action.

An analysis of the phenomenon of collaboration is usually predicated on the conceptual advances of Elinor Ostrom, who, starting in the 1960s, developed a canonical notion of collaboration linked with the challenges of managing common pool resources, or commons (Ostrom 2015). Ostrom argued that communities of collaborating humans could successfully realize governance of commons beyond the market's invisible hand or regulation and enforcement by the state. Despite this influence, the discussions and insights reflected in this volume show a need to move beyond this purely operational and instrumental three-pillar perspective on managing an identified resource. This raises the question of whether Ostrom's model is one specific use case in a much larger design space of possible collaborations. Collaboration exists and is also effective in the absence of a commons as a point of convergence in the real world. Collaboration is intrinsically and fundamentally future-oriented and, as such, a counterfactual state entertained by a collective. To understand and shape collaboration, we must thus include the ramifications of the narratives of present conduct and imagined futures that inform and guide or misguide a collaborating collective. Indeed, one could argue that ritualized play and games are forms of collaboration for the sake of collaboration itself, where the only constant is innate prosocial collective behaviors and the stories that shape specific collective action and interaction. Various chapters in this volume grapple with this broadening of the space of possible collaborations and open it to include mental processes and symbol systems.

Collaboration emerges from the interaction of goal-oriented human agents who voluntarily engage and communicate and are coupled by symbol systems expressing individual and collective memory, conventions, and norms. Phrased in these terms, collaboration occurs when all the complexities of human biology, psychology, sociology, culture, and embedding in the physical world are configured in specific ways. Each field of knowledge related to these levels of description is, at best, incomplete, let alone integrated with other fields. From this perspective, the phenomenon of collaboration is multidimensional, posing questions that require transdisciplinary treatment. This volume comprises contributions that both argue in favor of and against the traditional reductionist and instrumental perspective on human collaboration. Yet, the consensus is that effective research into its complexities must be developed across various disciplines.

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

This volume reports on a multiyear collaborative effort, convened by the Ernst Strüngmann Forum, to understand the nature and dynamics of collaboration and identify pathways for future research. New integrative programs are required to understand and shape human collaboration, building, expanding, and integrating existing efforts and their underlying theories, views, and practices. Looking beyond the horizon of current disciplinary boundaries suggests a need to develop a dedicated science of collaboration, also with an eye to the rapid intensification of hybrid human–machine collaboration and its intended and unintended consequences.

Contributors assumed different roles during the development of this project, at the in-person think tank, through written contributions as well as podcast interviews[1] (for further information, see the Preface). The initial in-person think tank was planned to take place in June 2021, yet the emergence of the COVID-19 pandemic made our gathering impossible. So we adapted—both in terms of scheduling and by expanding the framework created to pursue the topic. As such, this Ernst Strüngmann Forum itself served as a relevant case study of human collaborative behavior and contributed to an understanding of what facilitates and detracts from its success over time and in multiple settings.

The COVID-19 pandemic offered a further case study for analyzing how humans collaborate in the face of existential threats. Looking back, it is clear that despite myriad challenges, humanity demonstrated remarkable resilience and adaptability but also at immense costs of unnecessary suffering, loss of trust, and failing solidarity. Success stories abound, from the rapid adoption of virtual collaborative tools and new ways of organizing workplaces and education environments to the development of treatments and vaccines through global public–private team science efforts. Collective expressions of courage, empathy, dedication, and commitment ameliorated the effects of the pandemic and helped us as individuals, communities, and societies. Yet, it is difficult to look at this period as a triumph of human collaboration, especially if one takes into account the unprecedented magnitude of the financial incentives required to get the pharmaceutical industry to act, the lack of expediency and transparency in international data sharing to localize the origins of the virus affecting trust, and the dramatically uneven distribution of vaccines. Given that the next pandemic is merely a question of time, how can we get global health collaboration right next time?

Hybrid human–machine collaboration is a second challenge we are facing. Orchestrated by big tech, a new era of collaboration driven and modulated by artificial intelligence (AI) has emerged, the impact and dynamics of which are only partially understood. When OpenAI introduced ChatGPT to the public in November 2022, based on technology developed by Alphabet, it garnered one million users in five days. Since then, users have collaborated with large language models in domains traditionally believed to be the exclusive realm

[1] Podcasts are available at https://esforum.de/forums/ESF32_Collaboration.html?opm=1_3

of humans. A new technology-augmented form of collaboration is currently unfolding with unknown ramifications.

Our current situation has led members of the Bulletin of the Atomic Scientists to set the Doomsday Clock at 90 seconds before midnight, the closest it has been to the apocalypse since its inception in 1947. This unprecedented threat level reflects the challenges humanity faces, including nuclear escalation, armed conflicts, climate change, governmental and industrial malpractice, economic instability, resource scarcity, and novel biological, cyber, and technological threats. To respond to these challenges and mitigate escalating risks, large-scale collaboration aimed at the global common good is needed.

To address the complexities involved, the organizing committee for this Forum (Jenna Bednar, Julia Lupp, Bhavani Rao, Andreas Roepstorff, Paul Verschure, and Ferdinand von Siemens) defined a framework for the four discussion groups focused on the following core questions:

1. What is collaboration?
2. How do we collaborate?
3. What is the process by which we collaborate?
4. When does collaboration break down?

Given the multiple expertise involved at this Forum, the committee commissioned background papers to provide a starting point for the discussion. In addition, structured interviews with 27 people—experts who strive to realize collaboration in diverse real-life settings—contributed to the discussion. This volume synthesizes the observations, hypotheses, and tentative answers that emerged from this project. Interspersed between the chapters written for this Forum are summaries of these interviews (Chapters 6, 10, 15, and 19).

What Is Collaboration?

As mentioned above, we began with a broad definition: collaboration is understood as cooperation between agents toward mutually constructed goals. Four background chapters were commissioned to provide context to elaborate on this definition.

In Chapter 2, Stephanie Musgrave analyzes the evolutionary roots of collaboration. She distinguishes it from cooperation and sketches its underlying biological and psychological traits in great apes, or Hominidae. She identifies cognitive and motivational capabilities that underpin collaboration in experimental and naturalistic contexts: flexible coordination, shared intention, an understanding of causality, theory of mind, and roles. For instance, apes can recruit or solicit partners for joint actions, adopt different roles within a collaborative task, and use mutual communication to negotiate and coordinate actions. Apes can commit to joint goals and sustained efforts; they communicate with each other to achieve these and conform to social norms and group-specific behaviors that influence and stabilize collaboration. As they interact, apes demonstrate task and collaboration-oriented problem-solving skills as they adjust their actions to collaborative contexts. Collaboration in great apes is shaped and stabilized through value processes. Norms make conspecific behavior more predictable, whereas incentive structures shape it, including effort-proportional resource sharing, punishment, and rebuke of freeloaders. Learning and teaching behaviors (e.g., tool transfers) support skill transmission within the group and enhance a group's ability to collaborate. Previous experience with successful collaborations also influences future partnerships. Musgrave further addresses the foundational issue of the putative discontinuities in the cognitive and motivational underpinnings of collaboration in humans compared to other primates.

In Chapter 3, Chris Nierstrasz analyzes collaboration in trade and colonialism from a historical perspective. He observes that long-distance trade between Europe and Asia would not have been possible in the seventeenth and eighteenth centuries without extensive collaboration between and among Europeans and Asians. Partners entered and exited these collaborations expecting certain outcomes, which led to different results (from isolation to domination) depending on context. These outcomes were not always predictable and can be viewed as both the cause and result of the unfolding collaboration (for further discussion, see Chapter 5, 12, and 16). Nierstrasz argues that colonial control became possible when initial collaboration created codependency, leading to the construction of further mutually beneficial goals. Rather than use an implicit commons system to manage interactions (e.g., contract law or the concept of *Mare Liberum*) the collaboration process was stabilized through formalized written agreements between parties. This innovation was used to justify the actions of the Dutch East India Company (VOC) and reinforced their collaborative agreements with Asian rulers based on European notions of institutional stability and trust. Ultimately, the ability of the VOC to select collaborators and dictate outcomes paved the way for colonial exploitation (see also Chapter 6). Nierstrasz argues that historical collaboration, driven by practical and material needs, evolved into exploitative colonial actions, shaped by European and Asian opportunistic behaviors and institutional innovations. In analyzing the changing contemporary definitions of commons and global commons, Nierstrasz proposes that successful collaboration involves formalizing underlying agreements through written contracts or treaties, balancing adaptability and opportunism when adjustments are made to changing circumstances, and building stability through trust and consistent and reliable action.

In Chapter 4, Dana Dolghin and Chris Nierstrasz discuss a destructive aspect of the phenomenon by examining the relationship between the Dutch Jewish Council and German Occupation authorities during World War II. Collaboration between these two parties involved tactical compliance and a process of mutual, yet asymmetric, control based on inferring each other's intentions and values. Although it appeared as a choice, collaboration was more

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

an illusion as the Jewish Council was effectively coerced to engage with the German authorities, who dictated terms and exploited the Council's hope to influence outcomes amidst extreme power imbalances and the threat to its own survival. Collaboration served as a strategy to delay or soften the impact of the stronger party's demands. Dolghin and Nierstrasz examine the illusion of agency created by the Nazis and contrast it with the Council's belief that negotiations were possible. Its efforts to engage with the Germans, to increase their community's survival chances, were misguided and led to active participation in the deportation and murder of the Jews. Dolghin and Nierstrasz discuss the moral dilemmas and consequences of the Council's self-destructive collaboration. The Council's strategic rationale to mitigate Nazi policies was rooted in desperation, manipulation by the Nazis, institutional inertia, and misplaced beliefs in its own role and agency. The Jewish Council drifted away from its moral obligations by the path dependency of the coercive collaboration they had engaged in, which masqueraded as a genuine one. This chapter synthesizes the main properties and principles of destructive collaboration and defines themes for future investigation and analysis.

In Chapter 5, Ismael Freire and I illustrate the multidimensional nature of collaboration, encompassing biology, neuroscience, cognitive science, robotics and AI, engineering, social sciences, and the emerging field of collaborative AI. Using a thought experiment—a hypothetical Mars colony, *Dusk*, controlled and operated by algorithms, cyber-physical systems, and robots—we investigate what collaborative principles would be needed for this colony to survive (or fail) without human oversight. We highlight the need to understand the multiscale dynamics of collaborative processes, which requires multiple levels of description and explanation and propose that the construction of synthetic collaborative systems can help address methodological limitations of empirical research and constraints of experimentation. This will, in particular, assist in understanding the many feedback loops between the levels of organization within and between agents, their tasks, and their environments. Here, the emerging field of collaborative AI is relevant both in realizing hybrid collaboration between humans and machines and by building real-world proof-of-concept solutions to inform our understanding of collaboration. We advance the hypothesis that collaboration critically depends on the ability of agents to virtualize their environment, other agents, and self. This ability of virtualization allows agents to transcend the physicality of existence and interaction itself and be projected toward the future. We also introduce the notion of a morphospace, which points to the novel affordances that emerge from collaboration. This perspective opens up new avenues for the engineering of artificial collaborative systems. Synthetic collaboration will also permit the exploration of domains of collective action not explored through biological evolution, investigating open-ended evolution and adaptation in synthetic societies. The challenges and limitations of collaborative AI are highlighted, emphasizing the critical need to build synthetic and hybrid collaboration on integrated cognitive architectures and control in multiagent systems. We propose that answering these challenges defines the new field of collaborative cybernetics.

Closing out this section, Chapter 6 synthesizes insights from the podcast contributors, who were asked: What is collaboration? What is it good for? In describing their diverse experiences (from trade unions and critical health care to professional orchestras and bank mergers), various elements were identified as inherent to collaboration (e.g., cooperation toward mutual objectives, building of trust, a voluntary and affectionate union of parties). Surprisingly, practically none of the participants proffered a direct definition. All agreed that collaboration is a dynamic and complex phenomenon, not easily contained in a singular definition. It is an orchestrated effort integrating trust, shared vision, leadership, and proactive engagement to achieve collective aims. Collaboration transcends cultural, institutional, and disciplinary boundaries, embodying the essence of joint human endeavor and the pursuit or construction of a common purpose.

How Do We Collaborate?

To approach this question, three chapters were written to address how collaboration manifests itself in practice.

In Chapter 7, McLain et al. discuss how structured collaboration within self-help groups (SHGs) can result in sustainable economic and social benefits. They analyze the collaborative structure and benefits of community-level micro-financing through collective saving, credit systems, and social support mechanisms structured around the principles of mutual aid and accountability. The authors build their analysis on the Amrita Self Reliance Education and Employment (AmritaSREE) program, which spans 23 Indian states and involves about 13,000 SHGs. Created in response to the 2004 Asian tsunami, AmritaSREE initially sought to help women secure alternative income sources in affected coastal areas, and later in farming communities. Current activities now extend beyond direct finance to include education, life skills, and digital literacy for participants and their family members. Interviews with SHG members during the COVID pandemic identify the core principles of purpose, transparency, consensus, trust, education, within-group support, conflict resolution, and mutual support. McLain et al. also comment on potential risks (e.g., lack of transparency, reduced social cohesion, failing leadership) and ways to overcome these risks. They view the finances of an SHG as a common pool resource guided by design principles of resource management (Ostrom 2015).

In Chapter 8, Melody Ndzenyuiy and Heidi Keller discuss the impact of cultural context on collaborative practices, stressing heterogeneity in how collaboration is perceived and manifested across groups. They posit that the general understanding of human behavior is biased, as it is primarily informed by WEIRD (Western, highly educated, industrialized, rich, and democratic) knowledge bases—a bias that can misinform our understanding of collaboration. For instance, for the Nso' people of Cameroon, collaboration is a way of life, a notion that aligns with the belief systems of other subsistence cultures emphasizing community reliance over individual identity. This shared identity, or "we-ness," is deeply rooted in societies where communal bonds ensure mutual survival in a subsistence economy. From an early age, Nso' children learn to work together and serve the community in a self-initiated way. Extensive caregiving networks support their integration into the community. Through ethnographic analysis and in-depth interviews, they propose that the norms and values underlying this collaboration-centric organization have been key to its success. They discuss the differences that arise from polyadic and dyadic child-rearing styles and caution against applying WEIRD-derived methods and interventions onto community-based subsistence societies. The authors argue that valuing various cultural scripts and respecting context can help us achieve genuine collaboration between diverse groups.

In Chapter 9, Jônatas Manzolli and Julia Lupp provide a unique perspective on collaboration through the lens of art and performance. This setting demonstrates the core characteristics of collaboration as a dynamic, multiscale process between multiple agents who combine their capabilities and align their actions to pursue a mutually constructed goal. It also adds new elements imperative to successful collaborations. They demonstrate the relevance of these principles using three representative examples: a classical performance of Rachmaninoff's *Vocalise*, an interactive autonomous composing installation *Ada*, and an online collaborative artistic project, *Musical Letters*, realized during the COVID-19 pandemic. Further, they examine physiological and psychological aspects that underpin musical collaboration and suggest that trust, interpersonal relationships, and synchrony between agents play key roles, as does having a "common ground" and a shared understanding of musical elements and norms. They discuss feedback loops between agents (human and nonhuman), the phenomenon of "musical chills" or frisson, emotional engagement, and empathic relations implicit in musical collaborations. Examples of alignment highlight how cognitive and behavioral processes become synchronized: anticipating and adjusting to another's action may create aesthetic novelty. At a collective level, the authors argue that interpersonal relationships and power dynamics must be managed to ensure a successful performance. The chapter provides a framework for understanding the complex dynamics involved in musical collaboration, which can serve as a reference for understanding and shaping collaboration in other domains.

Chapter 10 synthesizes the experiences of podcast contributors, who were asked: What properties and traits do agents need to create and sustain collaboration? Their answers show that psychological factors shape collaboration in crucial ways: a shared understanding of common goals, shared belief systems, commonality of purpose and ideology, curiosity, enjoyment, trust, and transparent communication. They also provide practical examples where collaboration has been pivotal, shedding light on the mechanisms and outcomes of the underlying processes. These examples include the domains of health, resource management, religion, science, military, and philanthropy.

The Process

Four chapters were commissioned to address the focal question: How do we collaborate? Included in this section is a summary of insights provided by the podcast contributors.

In Chapter 11, Eric D. Beinhocker looks at a unique feature and foundational of human societies: large-scale collaborations between non-kin. He posits that relationships between individuals and collectives are a form of "social contracts." Thus, individuals must perceive a social contract to be fair for viable, effective large-scale collaboration to occur. Conversely, perceived unfairness can lead to withdrawal of engagement and even antisocial behaviors. Drawing on behavioral and social science literature, he proposes nine dimensions underpinning social contract fairness: agency, inclusion, dignity, rule-based governance, meritocracy, security, capabilities, reciprocity, and progress. Using these dimensions, he then analyzes the breakdown of political collaboration in the United States from the mid-1970s to the 2010s. During this period, special interests won out over the agency of individuals, and rule-based governance lost ground to undemocratic practices. Trust in the system was eroded, and perceived unfairness discouraged broad-based political collaboration. This period is also characterized by increased economic and social insecurities and stagnation: investment in the community decreased and personal and communal development became more unequal, further eroding the capacity for effective collaboration. Challenging standard neoliberal policies of the past, Beinhocker provides a comprehensive framework for understanding the psychological and social foundations of large-scale collaboration. He concludes with suggestions on how social contracts can be repaired and advances ideas for further research into social contract fairness.

In Chapter 12, Sander van der Leeuw explores collaboration based on theoretical insights and experience acquired from managing multidisciplinary teams with transdisciplinary goals, linking it to innovation, perspectives, and commons. Utilizing a complex adaptive systems perspective, developed in conjunction with the ARCHAEOMEDES project, he addresses diverse modes of collaboration across various scientific disciplines, their epistemological differences, and the challenges they face. He characterizes the collective nature of collaboration and competition in science (natural and life sciences and engineering) as being driven by shared, monothetic data and unifying frameworks and contrasts these with the more fragmented social sciences and humanities, which build on polythetic data. To overcome global challenges and sustain collaboration, van der Leeuw argues that we must recognize the integrated nature of the world: human activities and environmental changes are interdependent. Transdisciplinary constructive collaboration must be developed to address the resulting high-dimensional "wicked" problems. This implies a focus on the emergent scenarios of possible futures and their dynamic nature instead of the stability of what we call reality. In parallel to this characterization of the nature of global collaboration, van der Leeuw advocates for a transformative approach toward interdisciplinary collaboration to address global challenges, moving away from standard reductionist paradigms of Western natural science toward a paradigm where opposites are not contradictory but rather constructive complements.

In Chapter 13, Dennis Snower introduces the concept of recoupling. He argues that since both the scale and scope of humanity's collective challenges evolve over time, our collective capacities may become decoupled from these challenges. Human survival and flourishing depend on how successful we are in recoupling our capacities with our challenges. Such recoupling involves cooperation (working with others to achieve one's own goals) and collaboration (working with others toward common goals). When individuals collaborate, they participate in the purposes and welfare of the social groups in which they are embedded. He describes the building blocks of his recoupling thesis and analyzes how the scale and scope of our challenges can be aligned with the scale and scope of our capacities. As an example, Snower discusses global warming as a phenomenon resulting from the interactions among people and the emergent properties of Earth's climate system. Recoupling here requires a paradigm shift in measuring and pursuing success, moving away from purely economic metrics to include social and environmental indicators and integrating diverse disciplines and system thinking. Finally, he discusses internal and external mechanisms of collaboration, which need to work consistently in concert to achieve recoupling and outlines major implications for policy and business, in particular building on the notion of polycentric governance.

In Chapter 14, Scott E. Page explores the concept of collaborative architectures as they apply to informal groups and large, structured collaborations involving hundreds or thousands of participants, including those mediated by digital tools and AI. Architectures define who participates, how interactions are structured, roles and responsibilities, communication protocols, and how credit or blame is assigned. The success of a collaboration depends heavily on its architecture: it must fit the task at hand and be able to evolve over time as the collaboration changes. Page argues that the design space of the underlying architectures is vast, precluding a complete taxonomy. Thus, architecture design must be context specific and sensitive to cultures and their norms. One aspect that shapes collaboration is the structure of incentives. He discusses the need to adapt architectures to changing circumstances, based on iterative learning and feedback mechanisms that integrate internal psychological motives with external incentives and constraints. The sustainability of collaborative architectures requires the continuous reassessment and adaptation of these core aspects. Page poses questions about collaboration architectures' rigidity and the extent to which they should be centralized or decentralized. This opens the subsequent question of how architectures and their dynamics in roles and responsibilities can be implemented, scaled up, and maintained, creating a state of aporia.

Chapter 15 synthesizes the insights shared by podcast participants. The significance of trust, shared objectives, and navigating human complexities were underscored by all. Collaboration involves a commitment to continuous improvement, learning, and adaptation to address collective challenges and achieve common goals. As you peruse these varying accounts, you will note the need for a balance between structured rules and personal autonomy within collaborative efforts, integrating formal rules and informal norms, leveraging the strengths of structured governance and personal relationships. Effective collaboration requires transparent institutions, clear regulations, and the respect for individual autonomy of shared norms and values.

Discussion and Reflections

This section focuses on key questions that were posed to the working groups at the Forum: Why and how do humans collaborate, and what leads to a breakdown in collaboration?

Pacheco-Vega et al. (Chapter 16) address the central question of why agents collaborate as it applies to humans, animals, and machines. Their analysis includes the focal object around which collaborating agents coalesce. Here, the concept of the commons proved fundamental to the discussions (Ostrom 2015). The challenge, however, was to look beyond the commons. The authors analyze the conditions and processes that enable successful collective goal-oriented action. The "how"" of collaboration is elaborated by analyzing the psychological and social mechanisms that drive individuals to cooperate, including the vital role of shared goals and mutual benefits. They highlight the importance of trust and communication as foundational elements supporting the coordination of goal-oriented collective efforts. They posit that collaboration is most effective when adaptive; this allows participants to respond dynamically to changes and challenges. They put forth the idea that successful collaboration requires shared intentions and an adaptable framework that can evolve as the collaborative process unfolds and responds to new and unforeseen challenges. Such a framework acknowledges the interdependence between collaborating agents and is grounded in biological, social, and cultural values, norms, and conventions.

They also advance the psychological capabilities of collaborating agents, emphasizing the importance of acting relative to an imagined future or virtualization. The collaborating collective is guided by a shared imagination of what is possible in the future rather than what is actual in the present. In addition, Pacheco-Vega et al. critically examine the dominant economic models of collaboration and highlight their limited scope in fully capturing collaborative processes' multiscale, nonlinear nature. They argue that standard economic models typically undervalue collaboration's biological, social, and cultural underpinnings and often overlook how innate prosociality, shared narratives, cultural identities, and historical relationships influence collaborative dynamics, reducing these to transactional relationships and utility maximization. The authors call for a broader, more integrated, and holistic approach to understanding collaboration, especially now that it is expanding into the domain a hybrid collaboration between humans and machines.

In Chapter 17, Freire et al. explore how we collaborate by looking at its multidimensional nature, which is sensitive to various contexts, and the dynamism inherent in collaborative efforts. They stress the importance of understanding the cultural, social, and material conditions that impact collaboration. and emphasize that successful collaborations are context dependent. They warn that rigid, one-size-fits-all approaches often fail. Effective collaboration, instead, requires adaptive architectures that evolve over time, akin to complex adaptive systems. These architectures comprise various components (e.g., tasks, agents, and environment) that interact within a complex adaptive system, where internal dynamics and external feedback continually reshape the collaborative process.

Freire et al. question the necessity of fully aligned goals among collaborators and suggest that diverse and even conflicting individual goals can coexist within a well-designed collaborative framework. Success requires an effective collaborative architecture that supplies rules and roles that balance objectives while contributing to those of the group. This challenges the traditional belief that goal alignment is a prerequisite for successful collaboration. The merits of hierarchical versus flat structures are debated: the discussion that ensues focuses on the role of agency and power structures within collaborative processes, examining how individual and collective agency influence the dynamics of collaboration. Freire et al. hypothesize that understanding the interplay between agency and structure is crucial for designing more effective collaborative environments with a focus on power, flexibility, and adaptation. Freire et al. highlight the often-overlooked role of physical and cultural environments in shaping collaboration (e.g., the layout of a meeting room or the presence of fences), as this can significantly influence dynamics especially under resource constraints.

In Chapter 18, using historical and contemporary examples, DeDeo et al. examine whether there are design principles of collaboration that can be extracted to avoid future pitfalls and increase the chance of success. They posit that all collaboration breakdowns are fundamentally due to the misalignment of either values or actions. This insight challenges the conventional wisdom that successful collaborations require perfect alignment.

DeDeo et al. categorize these breakdowns into four types: catastrophic collapse, generative reboot, contested persistence, and sputter on launch.

Catastrophic collapse, such as the breakup of the Beatles, is attributed to value misalignment, whereas the collapse of the traditional *kapu* system (a set of laws and regulations in Hawaiʻian society) is linked to action misalignment engineered by elites. Generative reboots, like the transition from the Kyoto Protocol to the Paris Agreement, arise from value misalignments that pave the way for new collaborative frameworks. Contested persistence involves collaborations that endure despite internal conflicts, exemplified by the kinship tax in low-income countries, where individuals face conflicting demands from kinship obligations and market capitalism. Sputter on launch describes failed attempts to initiate collaboration despite agreement on goals, as seen in the global response to antibiotic resistance.

The authors suggest that human cognitive and cultural complexities make our collaborations more prone to collapse than other species (e.g., honeybees), whose evolutionary kin selection fosters more stable collaborations. In a practical sense, they advise policy makers to conduct an "alignment review" to assess how a policy change might affect the alignment of values and actions of the various collaborations within the society in question.

Chapter 19 summarizes insights into the breakdown of collaboration based on insights from the podcast contributors. Their responses highlight the disruptions caused by the lack of a shared sense of reality and truth, the mismatch of cultures, and failing leadership. The rise of social media and associated echo chambers has contributed to this. Leadership failures, miscommunication, lack of trust, and ethical breaches were identified as critical disruptors. Mitigation strategies from the interviews emphasize creating shared understanding and narratives, effective communication, trust, diverse team composition, and adaptive leadership. The COVID-19 pandemic, which was used as a test case, revealed the importance of global collaboration, the need for humility, and the role of leadership in fostering a sense of community and shared destiny. The resilience of collaboration is enhanced by including and valuing a mix of perspectives and skills, creating an inclusive space that values all participants' input, and promoting engagement using models like the "Blue Ocean Strategy." Processes must ensure all stakeholders are aligned in their objectives, avoiding competition-driven conflicts to maintain collaborative momentum. Leadership must be adaptable, balancing short-term and long-term goals while maintaining participatory governance.

In Chapter 20, I reflect on the complexities and future directions in collaboration research and practice, starting with overcoming the confusion in defining collaboration that this Forum revealed. I emphasize the evolving definition from its Latin origins into a complex and multifaceted concept that now includes diverse aspects beyond simple goal-oriented cooperation. Based on insights gathered during the Forum, I elaborate a research agenda to move the frontier of our understanding and practice of this complex phenomenon. First, the traditional focus on collaboration as a singular external goal-driven process and the generality of Elinor Ostrom's model on managing commons must be questioned. Based on these considerations, I advance an updated definition of collaboration, emphasizing its dynamics, intentionality, and future orientation. This broadens the agenda of our efforts to understand collaboration and includes the mental and symbolic processes involved instead of its operational aspects alone. Subsequently, I show that from this agential perspective, we encounter new challenges, paradoxes, and questions at the frontier of contemporary science. For instance, as soon as we include the first-person states of collaborating agents, we run afoul of the standard epistemological model of third-person Western science. This is more than a conceptual curiosity because it implies that to understand collaboration, we must redefine our scientific paradigm and move toward a multi-scale transdisciplinary framework that tolerates including first-person states such as experience and intentions. I also propose considering a continuum of collective dynamics from coordination to cooperation and collaboration, allowing various mixture states within a collaborating collective to exist and so move away from monolithic one-size-fits-all models. Linking collaboration to virtualizing imagined futures raises questions on its substrate and limitations. I explore the former by looking at the well-studied phenomenon of mental time travel, while the latter suggests a proximal zone of attainable collaborations predicted by the collective imagination. Taking the dynamics of collaboration seriously implies that it must be seen as an evolving and adversarial process; that is, a collaboration aims at realizing change in a task space that might not necessarily welcome this change. This implies that it is transient and transforms from exploration to exploitation once it succeeds, a transition that creates its own new challenges. Advancing the study of collaboration requires us to take a position on several foundational issues that have plagued the study of mind for millennia, such as what human nature is; the combined roles of phylogenetics, ontogenetics, and epigenetics; and the role of consciousness and free will. I identify a convergence in our view on collaboration with the notion of responsible autonomy underlying an understanding of free will. I also highlight the importance of diversity and dynamics in our understanding of collaboration using examples from attachment theory and stress. These examples point to the fact that each agent's control architecture is not static but dynamically reconfigures, dependent on task requirements. This process holds at multiple levels and illustrates that the dynamics of collaboration are continuously reshaped by feedback loops between the task, agents, and environment, necessitating a deeper understanding of collaboration's architecture. I also identify several paradoxes underlying the dynamics of collaboration, such as the tension between shared identity and schismogenesis, the adversarial nature of collaboration, the shift from creative exploratory collaboration to bureaucratic cooperation and coordination, and the associated risks of smothering collaborative dynamics. These considerations also reveal how fragile the collaboration process can be, dependent on individual agents' motivations, beliefs, and emotions. Lastly, I reflect on the future of collaboration as increasingly intertwined with technology, particularly AI, which presents both opportunities and risks for humanity and human collaboration.

From our initial, perhaps naive, definition of collaboration four years ago, we embarked on a journey that revealed collaboration as the most complex phenomenon in the universe in its multi-scale organization and the conceptual breadth required to study it. If ever a truly transdisciplinary and collaborative science was required, this is the moment for intellectual and practical reasons: our survival depends on it. The journey, facilitated by the Forum, has been instrumental in advancing our understanding of collaboration by allowing us to pose new questions and sketch the shape of possible answers. It has encouraged creative exploration of the problem space. In this, the Forum has truly achieved its goal of defining a roadmap for research that hopefully can make relevant contributions to addressing the pressing challenges we face in the epoch created by human collaboration: the Anthropocene.

With sincere gratitude, I wish to thank all collaborators in this project who participated in the Forum and the podcast interviews, my co-editors, the support of the Ernst Strungmann Forum, and above all, Julia Lupp's calm confidence and wisdom, without whom none of this would have happened.