

Collaboration

Interviewers: Paul Verschure (Convergent Science Network)
Jenna Bednar (University of Michigan)
Andreas Roepstorff (Aarhus University)

Welcome to the Ernst Strüngmann Forum podcasts—a series of discussions designed to explore how people collaborate under real-life settings. Joining us in the series are high-profile experts from diverse areas in society, whose experiences will lend insight to what collaboration is, what it requires, and why it might break down. This series is produced in collaboration with the Convergent Science Network.

- P. Verschure Hi, I am Paul Verschure and together with my colleagues Jenna Bednar and Andreas Roepstorff, we are speaking today with Robert Axelrod, the William D. Hamilton Distinguished University Professor Emeritus at the University of Michigan. Welcome Robert. Could you begin by giving us a sense of your background and where you today, as it relates to our discussion today?
- R. Axelrod I'm a political scientist at the University of Michigan. I've recently retired but am still active. I have long been interested in international security affairs and often use agent-based modeling and game theory to study and analyze such situations. I've also been interested in evolutionary biology, and I use evolutionary concepts in my political science work. Over the course of my work, I've engaged in many collaborations: one with the evolutionary biologist William Hamilton and more recently with an oncologist, Kenneth Pienta.
- P. Verschure What is collaboration? What is it good for?
- R. Axelrod In an academic setting, collaboration often takes the form of two or more scholars or students working on a single project with the hope that they can do something that none of them could have done alone. In many cases in my career, these have been interdisciplinary collaborations, where I've worked with somebody from outside political science or even outside the social sciences.
- P. Verschure What are the defining features of collaboration?
- R. Axelrod The basic defining feature is two or more people working on the same problem with mutual support, as opposed to independently and sewing it together at the end.
- P. Verschure One could argue that there are cooperative systems and collaborative systems, so why don't we just give them one label? In both cases, agents are doing things together. It seems that there are more nuances here at work.
- R. Axelrod I think collaborative efforts tend to be cooperative, so collaboration could be viewed as a subset that exploits the differences between the participants: the different things they bring to it, the diversity of talents and backgrounds and perspectives and goals. Cooperation is a much broader term. Cooperation could be, for example, that I trade something with you, and you give me money. Certainly, we usually think of collaboration as necessarily human, although you could say that ants collaborate as well.
- P. Verschure Could the difference be that collaboration has an explicitly defined goal, and in the case of cooperation there's an exchange interaction?
- R. Axelrod Collaboration typically involves a common purpose and the development of a common product. Something that they achieve together. It also typically involves specialization where people have different things that they bring to the team.
- P. Verschure With your background, interest in evolutionary biology, and identity as a political scientist, do you also see continuity in collaboration across different kinds of biological systems, from cells interacting together and differentiating in specific ways, as in your cancer example, to interacting conscious and intentional agents like humans?
- R. Axelrod I'll give you an example. Cancer involves cell lines developing the ability to overcome about seven or eight defenses that the host has to keep the cells in line. The conventional wisdom was that one cell line, based on a lot of mutations, would develop the ability to overcome all those defenses. I came up with an analogy that goes like this: Imagine two thieves trying to rob a

house. If one of them can pick the lock and the other one can turn off the alarm, they don't both have to have the capacity to overcome the defenses. That's a case of collaboration. It's also an example of cooperation. I have worked with geneticists and oncologists to develop this concept to see if it was feasible, and this has been found to be valid in lab experiments: there is collaboration within a tumor.

P. Verschure Defined in this way (i.e., in the cancer example), would you say that collaboration fits into the broader category of an emergent phenomenon, of interacting elements that create higher-level structure?

R. Axelrod Yes, absolutely. I was also going to add that that same example of tumors also represents a collaboration between three scientists—each of whom bring different expertise to the table. So, there is collaboration within the tumor as well as within the academic study of it.

P. Verschure The concept of the commons is often referred to when we speak about collaboration in humans; that is, there is some larger overarching objective that these interactive agents serve. Is there something like a commons in the case of cancer as a collaborative process?

R. Axelrod In the cancer case, I might say the common evolutionary path of the cells in a tumor is to multiply and increase, just as in every biological system. Some of these cells evoke something called VEGF [vascular endothelial growth factor], which encourages capillaries, which encourages blood supply, which encourages oxygen. So, one of them might promote oxygen in the immediate area while the other might promote some other nutrient. There's no common purpose defined by the evolutionary process itself, but the result could be the emergent property of a malignant tumor. In the case of typical human collaboration, there's often a specific goal that either has been handed to them as team members or they develop themselves. But individual human goals could be quite different: One person might be after fame, getting credit or promotions within an organization, while another might be motivated to try to solve a problem and care more about the problem than the personal credit. So, collaboration doesn't necessarily require complete agreement on the goals.

J. Bednar This is raising a lot of interesting questions for us, Bob. There are four "c" words that we're throwing around here: cooperation, collaboration, the commons, collusion (the existence of excludable goods). When we speak about a common interest or common purpose, it's interesting that we refer to two different systems: the human system (where there's rational forethought) and the biological system (where there isn't). We're talking about whether there needs to be completely overlapping common goals or whether there can be some differentiation in that. If the cancer cells are successful, they're actually going to kill the host, which will ultimately lead to their own deaths, so that's an interesting thing to unpack. At the human level, an interesting thing to unpack is the extent to which this interest held in common across all members of the group also includes benefits to those outside of the group.

R. Axelrod We've already talked a little bit about the first three. I see collusion as cooperation in the service of exploiting others. An example would be two major companies work together to set high prices; it's a collaborative restraint of trade. That kind of business collusion is done often in secret so that the outsiders who are being disadvantaged don't even necessarily know that the collusion is happening.

J. Bednar But is that what the cancer cells are doing then? Is it exploitation of others?

R. Axelrod Yes. With cancer, the evolutionary process, as you said, is a dead end.

P. Verschure A tragedy of the commons for the cancer cells.

R. Axelrod Yes, exactly. They're destroying the host that nurtures them.

J. Bednar The better they do, the more likely they're going to kill everything.

R. Axelrod Well, kill their environment. Fortunately, it's not infectious, so they don't kill others.

J. Bednar They kill themselves.

- R. Axelrod Yes, that's right. It reminds me a little bit of climate change.
- P. Verschure Andreas, do you want to step in?
- A. Roepstorff Your article "Achieving Cooperation under Anarchy: Strategies and Institutions" from 1985, really sets the groundwork for a lot of current thinking about collaboration and cooperation. In it you use the powers of game theory to outline dilemmas that could emerge. Could you discuss the difference between being in a prisoner's dilemma and another situation when you are mutually trying to collaborate with another person?
- R. Axelrod Collaboration itself could be a prisoner's dilemma. It could be that two members of the team each have an incentive to shirk, to let the other guy do all the work. But if they both shirk, then the collaboration doesn't accomplish anything. If the collaboration goes on for some time, either party might say to the other: "If you shirk, I'll shirk, but if you're willing to give your contribution, then I'll make a contribution, too." We can even play tit for tat: I'll start by giving some effort to this collaboration, and if you cooperate by giving effort as well, then I'll continue to do so. But if you start shirking and not contributing any effort, then I'll stop too. Then you'll realize that that was a losing proposition for both of us. Even within a collaboration, you could have a prisoner's dilemma between the participants.
- A. Roepstorff This suggests that mutual understanding of the situation is a critical part of collaboration.
- R. Axelrod People can have quite different conceptions of what it's about and what their goals are, but they have to appreciate that the process is interactive. So, what I do and what you do will jointly determine the payoffs to both of us. This must be understood. If you stop trying, you will come to realize that the other side might stop too, and that this is better than both of you working. But both agents could have quite different understandings of the situation. For example, you might have different goals and different understandings of what resources are available to the collaborators. You might have different understandings of who your audience is and what you're trying to accomplish. In many organizations, for example, two members of the same team might report to different bosses, and each one of them has the goal of satisfying their own boss. They don't necessarily have to have the same goals or even the same understanding of the situation. But they have this minimal need to appreciate that if they don't cooperate, the other side probably won't either.
- A. Roepstorff Critical to a game theory approach is the identification of a payoff matrix, but in many real-life situations, the payoff matrix is not clear cut. There is a lot of uncertainty. In your view, what sustains collaboration in situations where the payoff matrix might be unknown?
- R. Axelrod Typically, there might be a certain amount of uncertainty and error and noise in judging it, but it's rarely completely unknown. If it's completely unknown, then you'd have no reason to collaborate because you don't know whether it'll do you any good or not. I accept that there is some uncertainty about exactly what is common. Sometimes you're never even sure whether the other guy cooperated recently. That could be misperception and noise, but there has to be some understanding that your choices affect the other side and their choices affect you. The defining feature of the prisoner's dilemma is that the best payoff for you is if you are sure the other guy does all the work. The next best is that both of you work, and the next best is if neither of you work. The worst is if you do all the work, and the other guy shirks. That's all that you really need for the prisoner's dilemma: an understanding of the order of those possibilities. It's pretty common to do better by shirking, but you both do better by cooperating.
- P. Verschure Do you feel that game theory models, like the prisoner's dilemma, are still valid? They have been criticized for being caricatures of human decision making. Take the notion of bounded rationality. There's also increasing evidence that humans are not always optimizers. We optimize really for outcome. We might have surprising deviations from that. Do you feel that's still an effective model to look at the kinds of collaboration that we would like to optimize in everyday life and in our world?

- R. Axelrod Traditional game theory does assume rational actors who have foresight: if I do this, you do that. But evolutionary game theory has dropped the assumption of rationality. For example, cancer cells are not rational. They don't look ahead. They don't do any calculations, yet they still cooperate from an evolutionary perspective, which is to say that whatever works well means that they reproduce more. The prisoner's dilemma, for example, and game theory, in general, no longer require rationality. It's sometimes a useful assumption for making predictions. Sometimes the players aren't rational.
- P. Verschure There is a downside to that, because then you sacrifice goal-directedness, the coordination of human agents in collaborative contexts. This brings us to a behaviorist perspective on how human behavior comes about, which also has not been a complete success story. There is a price you pay if you bring it into this pure evolutionary, nonrational perspective, or do you feel that's a necessary step?
- R. Axelrod There's another version of the evolutionary aspect that is not rational: trial-and-error learning. You don't necessarily look ahead, but you experience whether you're doing well or not, and you change behavior if you're not doing well. Trial-and-error learning is another form of adaptation that humans can do and always do. That's different from rationality, but it's often a very effective thing.
- J. Bednar Another field to draw on for talking about human decision making is psychology. If people are not guided by clear incentives, but are still thinking, their thoughts can often take them wildly off the mark because people have biases, fears, etc. What role does that play? To what extent can we incorporate that into our game theory analyses? How can we make it work for us, not against us, as we're trying to understand human collaboration?
- R. Axelrod Let me give you a specific example. Three weeks ago, I was talking to Chinese and Russian academics, think-tank type people, in an ongoing discussion of cyber conflict and how to avoid instability. The point of what I said was to beware of vengeance. This is one of the psychological features of humans that Jenna was talking about, which is that when we feel insulted or exploited or disgraced or punished without reason, we often have a very powerful psychological need for vengeance, which we act on without much regard to the costs and benefits. We put all future-oriented thinking aside and say vengeance feels good and I need to do it. That's very dangerous. If you're in a realm like cyber conflict, when cyber weapons are used at scale, there's a lot of possibilities that people will feel the need for vengeance if attacked. The implication, of course, is that we should be aware of that when we try to understand the consequences of our own actions.
- J. Bednar But how so? Is it enough to say beware of vengeance, or is there guidance about how?
- R. Axelrod The guidance is to consider what the other side might regard as requiring vengeance. Here's a good example: Pearl Harbor. The Japanese attacked Pearl Harbor and they knew that the United States was stronger. But they didn't have a clear understanding of what would happen next, except that for six months they would be able to conquer everything in sight. After that they would be short on resources. The U.S. reaction to Pearl Harbor resulted, among other things, in a strong desire for vengeance. The Japanese, by contrast, figured maybe the Americans will let bygones be bygones; they'll take the Philippines and Indonesia and Malaysia and Indochina, but then we'll regard it as a fait accompli. But the U.S. didn't do this, of course, partly because of the desire for vengeance. The lesson is to be cognizant of what the other side (in this case, America for Japan) might regard as vengeful action. And if you decide to go ahead with it, you better do it with your eyes open.
- P. Verschure The wish for revenge is a very individually defined emotional state. How do we describe it at the level of a nation-state?
- R. Axelrod There it is shared by many individuals. I think it's perfectly reasonable to say that after Pearl Harbor, or 9/11, the U.S. was in a vengeful mood. After 9/11, we invaded Afghanistan and eventually killed Bin Laden. Part of the motivation was certainly vengeance.

- P. Verschure There are many layers, however, in between. We have institutions and larger groups of humans that all coalesce into something like a nation-state. How would that percolate between levels? How is vengefulness in the individual citizen incorporated in all the institutional layers?
- R. Axelrod Sometimes that can be very complicated. If you were trying to pass a law to slow collusion, you have to consider all those layers. In both examples I gave, it is straightforward: American citizens shared the experience of being attacked. They shared the experience of feeling that it was not an appropriate response to anything they did. The leaders felt the same way, and there was strong support in both cases at all levels. It didn't get complicated; you didn't have to worry about the role of Congress; you didn't have to worry about interest groups because there was such a widespread understanding of the situation as calling for vengeance. There was already identification with and loyalty to the nation. When Pearl Harbor was attacked, it wasn't just the Hawaiiansthat felt attacked, it was all Americans who felt attacked. Not just New Yorkers felt attacked on 9/11. There was a strong common understanding that the whole United States was attacked. To summarize, sometimes the situation is simple enough that we don't have to worry about all the intervening stages. They all operate smoothly.
- P. Verschure There's another aspect to this. Japan was involved in a systematic war of conquest since the 1930s. When they attacked Pearl Harbor, they were performing many military operations around the Pacific. Their goal was to knock the U.S. out of the war with one blow (they hoped all the ships were there). At the same time, the U.S. was competing for resources in the Pacific. An alternative interpretation could be that it was a competition for control and resources at the nation-state level, but to mobilize the citizenry to join the army and make sacrifices, vengeance was enhanced. In this way we could achieve the goals we have as a nation-state, which are not the same as vengeance.
- R. Axelrod I agree that goals involve resources and power politics. I'm saying that the motivation to declare war on Japan on December 7th—not a month or year before (from the American point of view, the situation hadn't changed dramatically in the weeks leading up to Pearl Harbor) was triggered by that event. What we did required leadership, including decisions such as how much of our American resources should be devoted to the European theater to fight the Nazis and how much should be devoted to Japan. That requires leadership and invites study of institutions to understand that choice.
- J. Bednar It sounds like there might be two different kinds of psychological processes going on. Maybe at the elite level, we've got a rational calculus of appropriate actions for the state to take, but, because it's a democracy, public support is needed so the elites can use emotion to get the public on board. Would you say that might happen?
- R. Axelrod That often happens. The clearest case may be the Spanish-American War and also the Mexican-American War, where leadership led the way for the public to support the goals of the war. It wasn't as spontaneous, from the bottom up, as it was on 9/11 or for Pearl Harbor. Certainly, leaders do try to mold public opinion. I remember Roosevelt had a famous line. He said to one of the people urging him to take action: "You've convinced me. Now go out and convince the public" so, they will force me to do it or support me doing it. There are certainly complex interactions between the public and the leadership. Once you get into that, you might well need to get into the institutional structures that mediate that difference, such as elections or the legislature.
- J. Bednar It may be that the public is unable to see as clearly as the elites the interests in acting or not acting, but then it's also probably the case that much information cannot be shared with the public. There's covert information that can't be released. Do you think that in our country we're thinking about the implications of the U.S. pulling out of Afghanistan? Do you think that this dynamic that you've been describing with Japan is happening in reverse in Afghanistan, in the sense that the public can't see our common interest in being present in Afghanistan anymore? And the leadership isn't able to make a convincing, compelling argument to bring the public along.

- R. Axelrod I don't think the leadership of the Biden administration sees value in continuing the war. I also don't think that they feel compelled by the public. I think they felt that they had quite a lot of choice about whether to declare a date by which they would withdraw. You made a comment to the effect that the elites are rational. I would say, not necessarily. They sometimes make choices that, after the fact, we think were made for bad reasons: not just bad information, but not carefully thought out. I've already mentioned the Japanese case. I thought that was an irrational decision. To make a comment about Jenna's statement that sometimes leaders can't share information, that's true. One of the factors that comes into play is trust. Does the public trust its leaders? A major loss in the last four years in the U.S. is a decline of trust in the government. For good reason.
- A. Roepstorff The notion of trust seems to be critical both when dealing with an uncertain situation or an uncertain payoff. It also seems to be critical in linking the individual level with the institution. How important is trust for collaboration, and how has the importance of trust, the understanding of the role of trust, changed over the years?
- R. Axelrod Collaboration requires trust—trust that the other person will give an honest effort to show up and do their part. Before you actually find out whether that happens, you might commit yourself to participate in this activity, trusting that it will be mutually valuable. In international affairs in the cyber context, the Russians and the Chinese have told the Americans that trust is perhaps the most important thing for making progress in avoiding war and escalation. Game theory has done some work on what is required to build trust, but I think there's a lot more to offer. Collaboration requires a certain amount of trust.
- A. Roepstorff Trust is also related to another critical issue, which is the temporal extension of the collaboration or the exchange, because trust involves projecting into a future. Is the relationship between future thinking and trust an area to develop further?
- R. Axelrod Absolutely. One of the major conclusions of my work on cooperation was the importance of what I call the shadow of the future: the idea that both sides know that there's going to be a continuing relationship. They don't discount the future too much in economic terms and view the future as worth worrying about, worth sacrificing for today by taking a chance on trust to see if you can get the long-term benefits. Trust can definitely be built up by an appreciation of that you're in for the long term.
- P. Verschure We have looked at several features, the last one being trust and the shadow of the future. If we want to put this all together in a comprehensive framework, these collaborative systems—from citizens through institutions to a president—must have some sort of architecture: different functions, protocols of communication, mechanisms in which incentives are distributed through the system, and so on. Can we look at collaborative processes from the perspective of an architecture? And if so, what would the main features of that architecture be?
- R. Axelrod One of the architecture's features is the extent to which the goals are set by an outside agent (e.g., the boss within an organization). When two members of a collaborative team are working for the same boss and have the goal of making the boss happy with their work, that's a pretty strong architecture and a pretty strong way of setting goals and monitoring them. Often, however, participants in the collaboration don't have a boss that sets the terms for the collaboration; they must work that out together, sometimes explicitly. For example, a treaty between two countries, a memorandum of understanding, or a contract between two companies (I'll provide this if you provide that). Sometimes it's formalized between the participants themselves. Often, though, it's less structured. There's a handshake, "let's work on this together," but no formal contract. There's not necessarily a detailed definition of what we're working on, or how much you'll do and how much I'll do. Sometimes in collaboration, there's an explicit statement that says I'll provide this if you'll provide that, based often on specialization. For example, when I collaborated with the Bill Hamilton to develop the biological implications of the Prisoner's Dilemma, we understood each other's strengths: he could provide lots of examples from biology for what I was talking about, and I could provide some of the

results of my computer tournaments and analyze them. We shared a common interest and understanding of both evolutionary theory and game theory. That helped us work together effectively. He knew a lot about beetles and I knew a lot about war. Those two knowledge spaces did not overlap.

- P. Verschure If I understand you right, you are saying there is a hierarchical structure of goals and subgoals. For the agents in the system to pursue those goals, we have something like trust, which implies that all these agents have a theory of mind. That's why you need trust. Because you have a model of the other agents and, just as they can do bad things, they can do good things. I assume that we'll do good things: that's trust, right? That already gives us elements of this architecture. You also have protocols of communication: both have to run within layers and between layers. Do you have a specific view on what kind of communication structure you need to maintain your goal structure?
- R. Axelrod It's often said that a major impediment to interdisciplinary collaboration is that the two sides don't speak the same language. When I started working on cancer, I certainly didn't speak the language of the cancer specialists, and there was a huge variety of specialized terms. You don't have to learn the 6,000 terms that are in the index of a major cancer textbook, but you do have to understand a lot. You have to learn a lot about what p53 is, for example, and why it's important, as well as VEGF and several hundred other things. Developing mutual understanding is often part of the collaborative process.
- A. Roepstorff So, architecture brings us back to scientific collaborations. A more anarchistic realm, so to speak, where there is not a clear hierarchy of a shared purpose or a shared goal and yet people figure out ways of collaborating. Should we think a little bit more about what it takes for these interdisciplinary collaborations to work? Is there still a commitment to some overarching shared goal, or shared cause, be it the advent of knowledge or the increase of knowledge, that stabilizes it? What does it take for agents in very separate systems to create this space in the middle between them? I think interdisciplinary collaborations are interesting instances of that.
- R. Axelrod My collaborations, which I don't think are typical, are more interdisciplinary than most. A project starts with the shared goal of achieving a better understanding of some phenomena in the outside world (e.g., an arms race or cancer). The intention initially, among academics, is that collaboration leads to concrete output in the form of at least one publication. Sometimes part of the goal is just having fun. I find it fun to work with people who know things that I don't know and to understand their perspective. A lot of science is driven by curiosity and a desire to play with ideas and concepts with the purpose of advancing our understanding and knowledge, but not necessarily a material purpose, such as getting a raise.
- A. Roepstorff You shared with us your presidential address to the American Political Science Association called "Political Science and Beyond." In it you discuss the four "ups" that researchers need to cultivate: reading up on a variety of fields, teaming up with others, loading up on research related to your problem, and lightening up when you need to escape from the problem for a while. Your principles could also be useful in initiating collaborations, as each assumes that you need to bring yourself into a particular state to allow collaboration to happen.
- R. Axelrod I call those things the four ups and they're on my website if people want to check out what they mean. The last one, to lighten up, is the most interesting. Whether you collaborate or work on your own, you can easily reach a dead end and you end up beating your head against the wall. It pays to lighten up, to back off and sleep on it. Turn to some other project for a while to get out of the current mode of thinking that you're stuck with.
- J. Bednar That makes me think about what causes a collaboration, which ought to continue, to fail. Or another way of thinking about it is, what makes a collaboration robust? What is it that causes something to fail? How could you structure it so that it would be able to continue despite being hindered by people being distracted by other projects?

- R. Axelrod I'll give you one example of a failure that I was involved with, which is a story about cooperation among cancer cells done with Dr. Kenneth Pienta. We sent it off to a Nature journal and received two reviews. Of course, what we were doing was somewhat speculative. One review stated: "It's impossible what they're proposing, this kind of cooperation." The other reviewer said: "Everybody knows this already." Obviously, this was a failure—a failure to get the article published in that journal as well as a failure in our ability to communicate what we were trying to do, to show that it wasn't impossible, and that it was different from what people already knew. The way we proceeded was to give up on that journal and take responsibility for explaining what's new here and why it was at least possible. It took a certain amount of perseverance to bounce back after receiving reviews that said that what we were proposing was impossible as well as already known.
- A. Roepstorff This notion of lightening up seems to be critically important for maintaining collaboration. Is there a payoff matrix for lightening up? How do we think of lightening up in terms of economic gain? It also opens a different dimension to understand trust by exploring what it means to make a long-term commitment to stay in the game together or creating a shared understanding. Does that have a payoff matrix in the traditional sense?
- R. Axelrod I think it has many payoffs, but at a discount rate. That is to say, how patient or impatient are you? How much does the future count, as opposed to getting immediate rewards? If the collaboration is to set up a business where you have one month to increase profits or else you fail, then you can't lighten up. You've got to work hard on that problem. But if you take a long-term perspective because your discount rate is too high, because you value the future, then you might come to the conclusion with your colleagues that what you're doing now isn't working. Then you ought to step back and examine why you structured the search for a solution the way you did; whether there are other ideas that are completely different from the ones that you've been trying to build on. I see it not so much as changing the payoff matrix as changing the discount rate, as well as realizing that to make progress you might have to give up on making progress in the short run, which wasn't working anyway, in order to explore distant possibilities that might pay off later.
- J. Bednar This makes me think about Jim March's excellent work on exploration versus exploitation. But there's another way of thinking about this and thinking about Andreas' question, which reminds me of the point you made earlier: the fun of working on a project is in part because you are working with other people. We don't typically think about this when we're doing game theory analyses. Generally, in game theory, we think in terms of extrinsic rewards, but there's also an intrinsic reward. Here I'm thinking particularly about the BACH Group, a group with which Bob has been involved for a long time—a beautiful collaboration between fellow scientists. I don't think it had a specific purpose other than being a place where you could share ideas in progress. You didn't have a definite goal beyond that in terms of a common publication. It was incredibly important to all of you who were participating in it, and a big part of that was the connection between fellow human beings. What role do you think this connectedness between humans plays in fostering successful collaboration and making it more robust?
- R. Axelrod You're quite right to emphasize the importance of human relationships, and for collaboration, the fun of working with other people. In work that has been done on political mobilization (e.g., working for a candidate, a party, or a cause), what's often found is that people get together at first to accomplish a particular purpose (e.g., women's rights or environmental sustainability) but then stay because of the personal rewards of working with like-minded people. It can often be an extrinsic motivation for a common cause, with the sustaining feature being that people enjoy working together. It's like a club. Often what helps to promote successful collaboration is when at least one of the members is explicitly intent on interpersonal relations, for example, by saying to somebody, "you contribute a lot, but why don't you listen to other people a little bit better?" I'm not very good at being tactful, so that's a bad way to put it. It has often been found in collaborative groups that one person takes on the role of a task leader focusing on the

task, and somebody else typically takes on the role of a solidarity leader, who is attentive to the interpersonal relationships and works to keep those on keel.

P. Verschure Returning to lightening up, it seems to be reminiscent of the idea of incubation which goes back to Helmholtz's theory of creativity, whereby you expose yourself to something followed by a period of distancing, of incubation, after which insight follows. This might be worth pursuing on some other occasion. However, I feel that conceptualizations and interpretations collide here. When we discussed the cancer example, we went into a model that denies intentionality and goal setting because the collaboration is purely a biological mechanical process. Yet when discussing the architecture of collaboration, the first thing you mentioned was goals. Could there be a gap between these two positions that shows that there's still a lot of work to be done to understand collaboration? What is the most critical question that we must answer to make progress on this topic?

R. Axelrod That's a big question. What is the most critical question...to deal with that question, we should lighten up and not give a 30-second answer right off the top of my head.

P. Verschure OK, although I was hoping for an answer.

R. Axelrod Sometimes you have a question that you're pondering that just bothers you. I had such a thing in graduate school when I was wondering, based on game theory, how you would measure whether a situation is like a zero-sum game or a partnership game where the interests are completely aligned and whatever helps one helps the other. In zero-sum, of course, whatever helps one hurts the other. How would you measure that? I couldn't do it. I put it aside, and several years later, I pondered it again. Then I woke up one morning and had, what I regarded as the answer, and I asked myself why it took so long. The answer, it turned out, is I had to forget something. One of the constraints I thought that you should put into this turned out to be irrelevant. You don't need it. After several years, when I came revisited the problem, I didn't bother with that constraint, which turned out to be very beneficial. So even when you have a question that's really bothering you, lightening up can be very helpful.

P. Verschure My last two questions will bring us over the finish line: Do you believe that humans will ever manage to maintain sustainable collaboration?

R. Axelrod Yes, in some regards. For example, we have sustainable collaboration about what the US dollar means. You can hand a person a US dollar any place in the world, and they'll give you an apple or whatever. There's an understanding, over just a piece of paper. Why should anybody take a green piece of paper in exchange for an apple? They are not going to eat the paper. The reason is that we all have a common understanding that somebody else could use that piece of paper to buy shoes or whatever. It's been the dominant world currency for almost 100 years, and it's likely to continue, even though it may not be the only major currency that's universally recognized. This is an example of sustainable understanding of a situation that promotes collaboration at a very high level. We need to have a sustained collaboration on climate change. That's difficult because everybody wants to burn coal because it's the cheapest source of energy today, but it's terrible. We can say to each other: "We'll try to reduce our coal if you try to reduce your coal." But there's no formal agreement or mutual monitoring or punishments for violation of the trust. There's a case where sustainable cooperation is absolutely essential, but it's really, really hard, partly because we're organized into nation-states.

P. Verschure If you could change one feature of humans, what would that one feature be?

R. Axelrod Do nothing. The reason I say that is if you change something, you don't know whether it will actually be an improvement, considering all the very difficult-to-understand side effects. You might think I'd like to eliminate disease. But if people live much, much longer, other considerations come into play. So, if I were to change one thing, I would do nothing.

P. Verschure Great. Bob Axelrod, thank you very much for this conversation.

R. Axelrod It was my pleasure. This collaboration with you guys is a lot of fun.