KOCHI UNIVERSITY OF TECHNOLOGY

Social Design Engineering Series *SDES-2020-11*

How can neuroscience contribute to the science of intergenerational sustainability?

Ryuta Aoki
Tokyo Metropolitan University

Ayahito Ito
Research Institute for Future Design, Kochi University of Technology

Keise Izuma
School of Economics and Management, Kochi University of Technology
Research Institute for Future Design, Kochi University of Technology

Tatsuyoshi Saijo
Research Institute for Humanity and Nature
Research Institute for Future Design, Kochi University of Technology

3rd August, 2020

School of Economics and Management
Research Institute for Future Design
Kochi University of Technology

How can neuroscience contribute to the science of intergenerational sustainability?

Ryuta Aoki[1*], Ayahito Ito[2], Keise Izuma[2,3], Tatsuyoshi Saijo[2,4]

[1]Graduate School of Humanities, Tokyo Metropolitan University, Tokyo, Japan
[2]Research Institute for Future Design, Kochi University of Technology, Kochi, Japan
[3]School of Economics and Management, Kochi University of Technology, Kochi, Japan
[4]Research Institute for Humanity and Nature, Kyoto, Japan

[*]Correspondence should be addressed to: R. Aoki (raoki@tmu.ac.jp)

Keywords: intergenerational sustainability, neuroscience, transdisciplinary approach

Author contributions: R.A., A.I., K.I., and T.S. designed the study. R.A., A.I., K.I., and T.S. wrote the manuscript.

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgments: This work was supported by a Grant-in-Aid for Scientific Research (A) to T. Saijo (17H00980).

Abstract

Intergenerational sustainability is an existential problem for humans, and coping with this issue requires large-scale cooperation extended across generations. However, recent empirical evidence suggests that people's concern for future generations is typically low, which is rooted from human's cognitive biases (e.g., temporal discounting and bounded empathy) and possibly exacerbated by modern social systems depreciating future generations' rights and voices. To achieve sustainable society, we need to design and implement novel social institutions that leverage our concern for future generations. In this paper, we discuss how neuroscience can tackle this fundamental challenge in collaboration with other disciplines. We review psychological factors and neural substrates that may underlie decision-making regarding intergenerational sustainability. We also propose empirical approaches to study neural bases of intergenerationally-sustainable decision-making. Notably, neuroimaging research has potential to unveil "hidden" neurobiological processes that are difficult to identify by behavioral observations alone. In addition, neural data can be used to predict real-world outcomes, which complements behavioral and self-report measures that may not always reflect true motives behind decisions. Understanding the neurocognitive mechanisms would provide insights into effective institutions that promote concern for future generations. We prospect that future neuroscience research will accumulate evidence from both laboratory and field experiments, thereby contributing to policy making and the transformation toward sustainable society.

1. Introduction

Intergenerational sustainability is an issue about the very existence of our species. Of particular importance on this topic is climate change, which is now considered as an existential threat for humans (Lenton et al., 2019). Environmental and ecological scientists have long been alerting to devastating effects of climate change on future generations (Rockström et al., 2009). Researchers in diverse fields (e.g., biology, economics, and philosophy), along with citizens, have been actively involved in collaborative actions to combat this issue. An interest in climate change has also been emerging in neuroscience (Aron, 2019; Aron et al., 2020; Langenbach et al., 2019). However, in the field of neuroscience, this is still sporadic movements by a small number of researchers. To date, there is no coherent trend of incorporating neuroscience into the transdisciplinary framework for sustainability science.

This article aims to discuss how neuroscience can contribute to solving this fundamentally challenging issue. The problem at the heart of intergenerational sustainability issues is the inescapable conflicts between the current and future generations (Saijo, 2015). We first argue that difficulties associated with intergenerational issues arise from several psychological factors. Next, we review neural substrates for these psychological factors, which may provide insight into human behavior regarding intergenerational sustainability. Third, we propose empirical approaches to examine brain processes supporting sustainable behavior. Lastly, we present open questions that should be addressed in future research.

2. Climate change as an issue of intergenerational sustainability

2-1. Why important: impacts on future generations

Climate change during the past 50 years is overwhelming (Steffen et al., 2007, 2015). This rapid acceleration in environmental changes is paralleled by explosive increases in socioeconomic indices such as global population and real GDP (Fig. 1A). Now we are living in an era called Anthropocene (Crutzen, 2002), where we ourselves substantially affect the Earth's ecosystems.

Climate change will induce global warming, sea level rise, and other changes in ecosystems (IPCC, 2014). It will also increase the risks of extreme weather events and outbreaks of deadly infectious diseases (Shuman, 2010; Stott, 2016; Wu et al., 2016). These changes may adversely impact future generations' life (e.g., industry, agriculture) and well-being. In fact, sea level rise has already affected citizens in island nations (e.g., Tuvalu), and recent extreme weather events (such as heatwaves, hurricanes, floods, droughts and wildfires) have caused huge damages in several countries. If climate change becomes worse, future generations will suffer more frequently and severely from these events. Accumulating evidence suggests that global climate change is man-made, meaning that our generation is critically responsible for well-being of future generations. Although the precise mechanisms of climate change remain unclear, human activity is the most plausible and parsimonious account, which is widely accepted among environmental scientists (IPCC, 2014).

The problem is that people often systematically underestimate the importance of intergenerational issues, in part due to human's cognitive biases. However, forecasts on economic impacts due to climate change may urge us to calibrate our perception. According to a recent report by International Labour Organization (Kjellstrom et al., 2019), the annual global economic cost of the productivity loss due to global heating (e.g., heat stress in daylight working) is expected to become ~2.4 trillion USD in 2030, which is a surge from 0.3 trillion USD in 1995. In addition, a recent estimation based on an economic growth model predict that the global GDP would reduce, if no effective measure on climate change is taken, by 2.5% in 2050 (roughly corresponding to US$ 5 trillion, assuming that global GDP in 2050 is US$ 200 trillion) and by 7.2% in 2100 (Kahn et al., 2019). If we consider long-term effects, economic damages induced by climate change are likely comparable to or even greater than those by mental disorders or by a pandemic (Bloom et al., 2011; World Bank, 2020). These estimates suggest that climate change will become a new plague for future generations, unless our generation takes immediate collective action.

2-2. Why challenging: psychological barriers

Despite the substantial importance, issues regarding intergenerational sustainability are fundamentally difficult to solve. Voices advocating immediate actions are rapidly glowing across the world, especially among younger generations. However, public opinions on climate change are divided, and attempts of international collaboration often fail (e.g., the US withdrawal from the Paris agreement). These facts show us that solving climate issues is quite challenging. Why is it so difficult? We discuss the possible reasons in below, focusing on psychological factors hampering our concern for future generation.

First, coping with intergenerational issues needs long-horizon goals spanning over a few generations (e.g., by year 2100). This inevitably causes temporal discounting (Frederick et al., 2002). If we assume a yearly discount rate of 3%, the subjective value of goods in 2050 is discounted to 41.2% relative to the value at present (in 2020). This means that an asset priced at $50 now is preferred over an asset priced at $100 thirty years later. Moreover, if one would not expect to live in 2100, she or he might consider the subjective value of goods in 2100 as zero (according to the *homo economicus* view). Second, humans have a limited ability to vividly imagine (or prospect) the far future. This perceived vagueness and psychological distance may reduce our naturally-occurring empathy towards future generations. Third, the inherent uncertainty of the future may elicit unrealistic optimisms (Sharot et al., 2011), offering excuses of not taking actions (e.g., “all problems will disappear by technological innovations”). These psychological factors may also be relevant to intragenerational issues, but likely be exaggerated in intergenerational issues.

Another critical factor is intergenerational conflict (Kamijo et al., 2017; Saijo, 2015), which has several common features with intergroup conflict (Fiske, 2002). Intergenerational issues arise from conflicts between the current and future generations (e.g., chasing economic growth versus pursuing sustainability). These conflicts may not have existed until recently (before the Industrial Revolution), but have become more stark as human activity reaches to the “planetary boundaries” (O’Neill et al., 2018; Rockström et al., 2009), i.e., the thresholds that define safe operating space for humanity on Earth (Fig. 1). If intergenerational conflict is overly emphasized, for instance by mass media messages, it may trigger antagonistic attitudes against future generations among a subset of people in the current generation. This response is irrational, in light of intergenerational sustainability, but probably congruent with the automatic tendency that humans often manifest in (intragenerational) intergroup contexts such as when competing over limited resources and soils. The antagonistic attitudes may further spill over towards members of the current generation who are in support of future generations, resulting in ideological polarizations within the current generation.

Importantly, the fact that psychological factors limit our concern for future generations suggests that better understandings of these factors offer clues for overcoming the limitations. Therefore, we expect that behavioral sciences (e.g., psychology, experimental economics, and neuroscience) for clarifying the mechanisms of human behavior will provide insights into solutions of intergenerational issues.

2-3. What we need: leveraging concern for future generations

So far we have argued that solving intergenerational issues is important but challenging, and that psychological barriers hamper our motivation to cooperate with future generations. What do we need to break the barriers?

Unfortunately, relying on “naturally-grown” empathy and self-control is an unlikely solution. Empathy is a key ingredient for altruism and prosocial behavior (Decety et al., 2016; Klimecki et al., 2016), but it might have evolved in small-size groups, and often is reduced for outgroups compared with ingroups (Cikara & Fiske, 2011). In addition, prosocial motivation (e.g., other-regarding preferences) decays as psychological distance increases (Strombach et al., 2015). Because we usually feel a greater psychological distance to the future proportional to its temporal distance, the spontaneous levels of empathy and prosocial behavior toward future generations are likely degraded. Self-control is the ability to resist temptations of sooner-but-smaller rewards and prioritize long-term benefits (Frederick et al., 2002). However, behavioral economic studies have clearly shown that humans on average have limited levels of self-control (Thaler & Benartzi, 2004), often resulting in suboptimal choices in term of intertemporal rationality in everyday situations (e.g., dietary choices, pension plans). Critically, dilemmas involved in intergenerational-sustainability issues are often much harder to resolve compared with those involved in personal intertemporal choices, because delay periods are much longer (e.g., over generations) and recipients of long-term benefits (i.e., future generations) are not those who exert patience (i.e., the current generations). Because of these reasons, the empathy we naturally feel to future generations, as well as the self-control we exert for future generations, might be limited.

Consequently, the most important goal for achieving intergenerationally-sustainable society is to develop and implement novel social systems (or “institutions”) that effectively leverage the current generation’s concern for future generations. Otherwise the current generation would overexploit intergenerational common pools and put future generations into a crisis. The role of neuroscience, together with other fields in behavioral sciences, is to accumulate empirical evidence for effects and mechanisms (e.g., biological underpinnings) of such institutions, thereby contributing to policy making for sustainable societies. It is obvious, however, that the entire issues of sustainability cannot be solved by neuroscience alone. We need a transdisciplinary framework that facilitates close collaboration among diverse research fields as well as with citizens and policy makers.

3. Insights from existing neuroscience studies

What can neuroscience exactly do? Identifying brain regions involved in intergenerationally-sustainable behavior would be a starting point (if not a goal). To date, neural substrates for intergenerationally-sustainable behavior remain poorly understood.

However, findings from existing neuroscience research may provide useful insights. Here we review past neuroimaging (e.g., functional magnetic resonance imaging [fMRI]) studies on related topics. We particularly emphasize that neuroimaging studies, if combined with appropriate experimental designs, may offer useful insights into psychological processes that cannot be obtained by behavioral observations alone.

3-1. Prospecting the future

The ability to think about future generations relies on our ability to vividly imagine and simulate the future with episodic details. This cognitive function is referred to as “prospection” or “episodic future thinking” (Gilbert & Wilson, 2007; Schacter et al., 2017), and considered as unique to humans (Suddendorf et al., 2018). Previous neuroimaging studies have shown that brain regions such as the medial prefrontal cortex (mPFC), precuneus, and temporoparietal junction (TPJ) are involved in prospection, with especially important roles of the anterior PFC (aPFC; also called the frontopolar cortex) in representing future goals (Brown et al., 2016; Doll et al., 2015). These brain regions are overlapping with the so-called default-mode networks (DMN). Interestingly, the DMN is also implicated in creative thinking (Beaty et al., 2018), probably because both prospection and creative thinking requires counterfactual thinking and imagination (Hassabis et al., 2007).

3-2. Prosocial behavior

Intergenerationally-sustainable behavior is by nature prosocial, because it benefits future generations while (typically) imposes costs on the current generation (Langenbach et al., 2019; Saijo, 2015). Neuroimaging studies over the past two decades have revealed sets of brain regions involved in prosocial behavior (J. K. Rilling & Sanfey, 2011; Ruff & Fehr, 2014). Intergenerationally-sustainable behavior is likely mediated by similar brain regions, although this should be empirically tested. These brain regions include those involved in value-based decision-making (i.e., decisions made on the basis of subjective value—such as when you consider which of “receiving $10 now” or “receiving $20 after a month” you prefer), and those involved in social cognition, such as the ability of inferring others’ mental states (Fig. 2).

An important issue is whether certain treatments/interventions can leverage prosocial behavior within individuals, so as to know if a given institution can promote prosocial behavior at the societal (collective) level. Past neuroimaging studies have shown that prosocial behavior can be enhanced via multiple distinct neural pathways. A key implication here is that, even if two treatments yield behaviorally similar effects, the neural mechanisms underlying these effects could be distinct. For instance, Hein et al. (2016) showed that experimental treatments inducing “empathy-driven altruism” and “reciprocity-driven altruism” promote prosocial decisions (i.e., giving money to others at the cost of self-interest in a laboratory decision-making task) to an equivalent extent, but their effects are mediated by different brain mechanisms. In their fMRI study, “empathy-driven altruism” was operationally defined as increased prosocial decisions after observing another person who received painful electrical shocks (“I help you because you are suffering”). On the other hand, “reciprocity-based altruism” was defined as increased prosocial decisions after observing another person who did a kind act to the study participant (“I help you because you helped me”). The two treatments elicited distinct patterns of fMRI signals among key brain regions supporting prosocial decisions (the ventral striatum, anterior insula, and dorsal anterior cingulate cortex).

This finding gives an exemplar case where neuroimaging can distinguish different psychological motives (i.e., empathy and reciprocity) behind behaviorally indistinguishable prosocial decisions.

It is worth noting that the terms “empathy” and “reciprocity” are multi-facet concepts. Social psychology and neuroscience have investigated how these concepts are comprised of distinct factors, each of which may have different effects on behavior. Empathy can be divided into “affective empathy” and “cognitive empathy,” which are subserved by distinct brain systems (Shamay-Tsoory et al., 2009). Affective empathy is the ability to share emotion with others (e.g., “emotional contagion”). This function is often automatic and accompanied by visceral responses (Decety et al., 2016). For instance, when we observe another person who is wounded and bleeding, we spontaneously feel the pain that the person would feel. Cognitive empathy is our ability to infer others’ mental states (e.g., intentions and beliefs), which is also referred to as “theory of mind (ToM)” or “mentalizing” (Frith & Frith, 2003). ToM is critical for cooperative behavior (by allowing us to share intentions with others) from hunting in human ancestry societies to resolving international conflicts in the modern world, while it is also critical for strategic behavior such as bargaining (by enabling us to predict and outsmart others’ intentions).

Likewise, reciprocity can be divided into several distinct concepts. One basic distinction is between direct reciprocity and indirect reciprocity, with the latter considered to be indispensable for large-scale cooperation among genetically unrelated individuals (Rand & Nowak, 2013). Indirect reciprocity can be further divided into upstream (or “pay-it-forward”) reciprocity and downstream (or “reputation-based”) reciprocity. An fMRI study showed that upstream and downstream reciprocities have different neural substrates (Watanabe et al., 2014), which is another exemplar case where neuroimaging provides evidence for dissociations between behaviorally similar concepts. It remains unclear whether the two types of indirect reciprocity are relevant to intergenerationally-sustainable behavior. Intuitively, only upstream reciprocity would contribute to intergenerationally-sustainable behavior, as downstream (reputation-based) reciprocity do not work between (non-overlapping) generations because of the asymmetry of time. An intriguing open question is whether an extended version of reputational concerns, such as motivations for leaving a legacy (Wade-Benzoni et al., 2010), plays roles in facilitating intergenerationally-sustainable decisions.

Diverse concepts and sub-concepts are relevant to human prosocial behavior, but they may have different effects on intergenerationally-sustainable behavior. Careful distinctions between these concepts are especially important in transdisciplinary research, because misuse of the terminology may result in confusion. For example, one may want to claim that “empathy promotes intergenerational cooperation.” However, a certain kind of empathy may increase cooperation within a small group but may simultaneously enhance aggression toward outgroups (Bernhard et al., 2006; Bruneau et al., 2017). If this is the case, this type of empathy (i.e., parochial empathy) may not be beneficial for (or even backfire) large-scale cooperation required for intergenerational sustainability. A more rigorous behavioral and neuroscientific research for clarifying these inter-related concepts to avoid such confusions.

Neuroimaging could be useful to predict prosocial behavior especially when self-report measures are not reliable predictors of actual behavior. There are several reasons that self-report measures do not necessarily tap into true motives behind behavior. Social decisions and evaluations are often influenced by implicit brain processes (Stolier & Freeman, 2016). If an individual is not aware of these implicit processes, she or he may not be able to report the true motives behind decisions. In addition, self-report measures are susceptible to reporting biases such as demand characteristics. For instance, self-reported intentions of engaging in real-world sustainable behavior (e.g., using carpools) may not reflect actual behavior (Kristal & Whillans, 2020). In such cases, neural data may outperform self-report measures in predicting actual prosocial behavior.

3-3. Intertemporal decision-making

Neuroimaging studies have revealed brain mechanisms involved in intertemporal choice (Kable & Glimcher, 2007; McClure, 2004), although these studies mostly deal with decisions concerning self alone (e.g., trade-offs between the present self and the future self, with no relevance to other persons). A key brain region is the dorsolateral prefrontal cortex (dlPFC), a region important for self-control in decision-making (Fig. 2). Experimentally modulating dlPFC activity by brain stimulation techniques (such as transcranial magnetic or current stimulations) influence temporal discounting, such that diminished dlPFC activity makes individuals more impulsive (Figner et al., 2010). Another brain region implicated in intertemporal decision-making is the aPFC, a region important for metacognition and prospection (Fleming et al., 2010; Gilbert & Wilson, 2007). An fMRI study showed that the aPFC is activated when people are aware of the temptation of sooner-but-smaller rewards and precommit to restrict the access to the tempting options (Crockett et al., 2013). Another line of studies has shown that episodic future thinking (e.g., prompting people vividly imagine future events) can decrease temporal discounting (Peters & Büchel, 2010), possibly by reducing perceived distance and ambiguity of the future. These mechanisms relying on self-control, metacognition/prospection, and imagination may serve as distinct (but inter-related) pathways by which people can make intertemporal rational decisions (Bulley & Schacter, 2020).

As noted above, these findings are derived from studies that examine intertemporal decisions involving rewards for self alone. It remains to be addressed whether any of these mechanisms promote intergenerationally rational choice, which involved tradeoffs between the present self and the future others. For instance, it could be the case that self-control alone is not sufficient for intergenerationally sustainable decisions (Langenbach et al., 2019), but stimulating people's imagination about the future opens the gate for self-control to exert effects on sustainable decisions (i.e., an interaction between imagination and self-control). This hypothesis is in line with the hierarchical organization of the brain, where the aPFC situates the upstream of the dlPFC (Koechlin, 2003).

3-4. Intergroup conflict

Although how the brain react to intergenerational conflict remains unknown, neuroimaging research have studied the mechanisms underlying intergroup conflict. These studies typically focus on racial/ethnic groups (or supporters of different political/sport teams), and examine neural correlates of intergroup behavior (e.g., ingroup favoritism and outgroup hate). Brain regions implicated in automatic emotional responses (e.g., the amygdala), social cognition (e.g., dmPFC), and affective judgment (e.g., vmPFC) underpin negative attitudes toward outgroups such as prejudice and discrimination (Amodio, 2014). These neural substrates may provide mechanistic explanations of why people tend to fear outgroups (often automatically) and unfavorably evaluate them (e.g., less trustworthy, less competent). For example, the amygdala is known to play a key role in fear learning (an association between neutral and aversive stimuli), and fear learning is considered as an underlying mechanism of how prejudice toward an outgroup is acquired in a real world (Phelps et al., 2000). Consistent with this idea, we recently showed that activation patterns in the left amygdala were significantly associated with implicit evaluations (i.e., prejudice) toward an ethnic outgroup (Izuma et al., 2019). Notably, behavioral and neural biases against outgroups emerge even with experimentally created groups (e.g., by minimal group procedures). This may raise the possibility that excessively emphasizing the border between the current and future generations elicits negative attitudes (e.g., decreased empathic care and neglect) toward future generations. Instead, messages emphasizing continuity between the current and future generations (e.g., by emphasizing that the act of our generation will be bequeathed to future generations as legacies) may reduce the tensions between generations. This idea could be tested in an experiment that contrasts treatments emphasizing competition (vs. continuity) between generations.

Intergroup conflicts between races and ethnicities in the real world seem to be harsh and robust. Can we reduce it by interventions? An fMRI study showed that receiving helps from outgroup members promotes empathy toward the outgroup, which possibly relieves intergroup conflicts (Hein, Engelmann, et al., 2016). This effect was underpinned by a prediction error signal (in a reinforcement learning process) observed in the anterior insula, consistent with the region’s role in affective empathy. Similar processes likely occur in the real world, for example when immigrants interact with local residents. However, the same process may not work in intergenerational situations, because the current generation has no opportunity to be helped by (remote) future generations. Reducing intergenerational conflicts may thus be more difficult compared with reducing intergroup conflicts within a generation, and necessitate novel methods of interventions.

3-5. Connecting the dots

It is important to note that brain regions work together as distributed systems. For instance, previous work has shown that the TPJ exhibits "functional connectivity" (a term indicating statistical dependence of fMRI time series between distant brain regions) with regions involved in value-based decision-making such as the striatum and vmPFC (Park et al., 2017), and this functional connectivity underpins prosocial motivation called warm glow (Andreoni, 1990). Functional orchestration among different brain regions is critically important for decision making. For instance, the vmPFC shows functional connectivity with regions such as the TPJ, dlPFC, and aPFC in context-dependent manners during decision making tasks (Baumgartner et al., 2011; De Martino et al., 2013; Hill et al., 2017). To better understand neural bases of intergenerationally-sustainable decisions, we need to examine how multiple brain regions work in concert as distributed brain networks.

4. Possible empirical approaches

How can we empirically study neural bases supporting intergenerationally-sustainable behavior? Here we propose two types of approaches that would be useful in future research.

4-1. Using economic games: decision-making in laboratory settings

The first approach is to use behavioral economic games designed to study integenerationally-sustainable decision-making, and incorporate them into neuroimaging experiments. In this approach, participants make decisions regarding intergenerational sustainability while their brain activity is measured using neuroimaging techniques. Previous neuroimaging studies have used similar approaches to examine neural bases of social decision-making. For instance, economic games such as the trust game, prisoner’s dilemma game, and ultimatum game have been used to study trust, cooperation, and inequality aversion, respectively (McCabe et al., 2001; J. Rilling et al., 2002; Sanfey et al., 2003).

The advantage of this approach is that it can examine neural responses in well-controlled laboratory settings. Behavioral economic games allow researchers to systematically manipulate experimental variables, such as the cost and efficiency of sustainable decisions, as well as to examine effects of certain treatments versus well-matched control conditions. This is particularly useful when combined with computational modeling (Behrens et al., 2009), which enables to decompose a decision process into distinct components and identify how experimental manipulations (e.g., a treatment) affect each component.

A few behavioral economic games suitable to study intergenerational sustainability have been proposed (Fischer et al., 2004; Hauser et al., 2014; Kamijo et al., 2017; Langenbach et al., 2019; Sherstyuk et al., 2016). In these games, a group of players (or an individual player in some studies) represents a “generation,” and decisions are made successively from one generation to another (Figure 3A). Of note, a decision made by a generation affects only later generations, but not the other way around. This mirrors the unidirectional nature of intergenerational dependencies between the past and future generations in the real world (i.e., the time flows only from the past to the future).

The essential characteristic of these games is that they embed intergenerational sustainability dilemmas (i.e., trade-offs between the current and future generations' benefits) into the game structure. For instance, in the Intergenerational Goods Game used by Hauser et al. (2014), each generation consisting of five players makes a collective decision (on the basis of median voting) as to how much they extract a resource from an intergenerational common pool. If the extraction by the current generation is under a predetermined threshold, the pool will be replenished and the next generation will play the game in the same way as the current generation does. On the other hand, if the extraction level exceeds the threshold, the pool will be exhausted and the following generations will lose the opportunity to play the game. Thus, an overexploitation of the intergenerational common pool benefits the current generation, but harms the future generations. In the Intergenerational Sustainability Dilemma Game (ISDG) used by Kamijo et al. (2017), each generation consisting of three players makes a collective decision (on the basis of conversations among the players within each generation) between “Option A” and “Option B” (corresponding to “unsustainable” and “sustainable” options, respectively). Although the current generation receives a larger payoff by choosing Option A compared with Option B (say, $36 vs. $27), if the current generation chooses Option A, the next generation will face a similar decision between Option A and Option B but with reduced payoffs for both options (e.g., $27 vs. $18). On the other hand, if the current generation chooses Option B, the next generation will face a binary decision with the payoffs maintained for both options (i.e., $36 vs. $27). This means that the pool gradually decreases and eventually is depleted if many generations choose Option A, whereas it is sustainable as long as all generations choose Option B. Thus, each generation face at a dilemma between self-interest (i.e., choosing Option A) and sustainability (i.e., choosing Option B). The game used by Langenbach et al. (2019) also entails a similar feature of intergenerational sustainability dilemma (i.e., forgoing self-interest to achieve intergenerational sustainability), although the specific implementation differs from the other games.

There are some important variations among the games used in the previous studies (Figure 3B). These games typically implement intergenerational dependencies as Markov processes (i.e., the payoff structure of generation t+1 is solely determined by the decision of generation t, irrespective of the decisions of generations t–1, t–2, …), and the current generation is not provided with the information about past generations. Therefore, the players' decisions would be influenced by "prospective" factors (i.e., how their decisions affect the next generation), but not by "retrospective" factors (i.e., how the past generations have made decisions). This simplifies the structure of the games, but misses the important aspect of the intergenerational decisions in the real world—that is, the decisions of the current generation are influenced by the history made by the past generations. A notable exception is the ISDG (Kamijo et al., 2017). In the ISDG, each generation is provided with the full history of the past generations' decisions. This allows researchers to examine effects of retrospective factors (i.e., the history) in addition to those of prospective factors. For instance, a player's decision may conform to past generations' decisions, because the choices made by the past generations may serve as a reference point or set a "norm" (Xiang et al., 2013).

Alternatively, some players may want to "break off bad habits" and behave benevolently to the future generations if several past generations have consecutively chosen unsustainable options. Although the effects of past generations' decisions on the current generation's decision could be complicated, investigating such retrospective factors would provide important insights into understanding intergenerational decisions in the real world.

Another unique aspect of Kamijo et al. (2017) is that it allowed conversations among the players within each generation (but not between generations) before they make a collective decision. This is unlike the studies done by Hauser et al. (2014) or Langenbach et al. (2019), where participants were prohibited to make communications. Conversations and communications are indispensable parts of policy-making processes in the real world (e.g., deliberative democracy and procedural justice). However, their effects on collective decisions remain unclear, for instance whether communications among individuals lead to collective wisdom (Bahrami et al., 2010; Navajas et al., 2018) or induce phenomena such as risky shift and group polarization (Lord et al., 1979). In particular, it is possible that conversations alone cannot facilitate intergenerationally-sustainable decisions without additional institutions (Timilsina et al., 2017). Thus, how conversations and communications affect collective decisions regarding intergenerational sustainability is worth investigating. Letting participants freely converse with others in the MRI scanner involves technical difficulties, because it induces head motions and decreases fMRI signal quality (if not impossible; e.g., Chen et al., 2017). Instead, experimenters can let participants converse with others under certain conditions (which serves as experimental treatments) outside the scanner, and then let them perform decision-making tasks (without conversations) in the scanner.

This allows researchers to examine how neural responses during the decision-making task is affected by prior experience of conversations (under a certain condition).

A major concern of the approach using economic games is whether behavior observed in laboratory experiments translates to real-world sustainable behavior (from purchasing green products and using reusable bags to expressing support for pro-environmental policies). Because the games to study intergenerational sustainability are developed relatively recently, research ensuring ecological validity is still lacking or scarce. For other economic games widely used in past research (e.g., the dictator game, ultimatum game, and trust game), intensive efforts have been made to ensure their ecological validity—that is, decisions in these games reflect real-world behavior regarding generosity, fairness, and trust (Franzen & Pointner, 2013). Similar efforts are needed to establish the correspondence between behavior in the laboratory games concerning intergenerational sustainability and real-world sustainable behavior.

We also emphasize that a central goal of research is to develop institutions that are effective in real-world situations. Some experimental treatments used in laboratory settings may not work in the real world. For instance, Hauser et al. (2014) showed that median voting is effective in sustaining intergenerational common pools in a laboratory setting. However, median voting may not work if the majority of voters prefer self-interest over sustainability and the median exceeds a limit of sustainable resource provision. In fact, when we look around the real world, most countries are not even close to heading toward sustainable societies (Wackernagel et al., 2017). As a result, agreements reached in international conferences such as the Conference of the Parties to the United Nations Framework Convention on Climate Change (COP) and the World Economic Forum (WEF) are often not sufficient to achieve sustainability (Steffen et al., 2018), and criticized as putting future generations still under risk. This may suggest that different types of institutions are needed to uplift concerns for future generations.

4-2. Using naturalistic stimuli: predicting real-world outcomes

The second approach focuses more on predicting real-world outcomes from brain data (Fig. 4). This approach uses naturalistic stimuli, such as movies and media articles that deliver persuasive messages promoting certain sustainable behaviors (e.g., purchasing green products, using carpools, and reducing air travel). Participants are presented with naturalistic stimuli inside the scanner, like when they watch TV commercials or read new articles in everyday life. The aim of research is to associate neural responses with behavioral/attitude changes induced by persuasive messages. Similar approaches have been used to predict behavioral changes induced by health messages (e.g., quitting smoking, using sunscreens) from neural responses (Chua et al., 2011; Falk et al., 2010). Recent advances in spatiotemporal analysis of fMRI signals have enhanced the utility of naturalistic stimuli. For instance, voxelwise encoding models allow to investigate neural representations of various low-level perceptual (e.g., audiovisual) features and semantic contents included in naturalistic stimuli (Huth et al., 2016). In addition, analyses looking at brain synchrony among individuals (e.g., inter-subject correlations of fMRI time series) allow to capture temporal dynamics of neural responses in data-driven manners, without requiring pre-specified stimulus onsets (Sonkusare et al., 2019).

This approach can be used in two distinct ways. The first way is to predict individual differences in behavioral changes. The same message may induce different degrees of behavioral changes across individuals, and the inter-individual variations might be associated with neural responses in specific brain regions. Previous studies have shown that activation in brain regions processing personal relevance of stimuli (e.g., the mPFC and precuneus) is associated with individual differences in behavioral changes (Falk et al., 2010). Using brain data to predict behavioral changes is particularly useful when self-reported intentions of changing behavior are biased by confounding factors (e.g., social desirability) and not reliable predictors of actual behavioral changes.

The second way is to predict collective outcomes in the real world using brain data obtained by laboratory neuroimaging experiments. In other words, using neural responses observed in “neural focus group” to forecast population-level behavior in large-scale social groups (Falk et al., 2012). Previous studies have shown that brain data obtained from small sample-size groups (e.g., around 40 participants) can predict population-level outcomes such as viral sharing of news articles on social networking services and aggregate view frequency of YouTube videos (Scholz et al., 2017; Tong et al., 2020). These studies suggest that activations in brain regions implicated in processing of personal relevance (e.g., the mPFC) and reward values (e.g., the striatum) are predictive of population-level outcomes. Such approaches may also be useful to examine effects of institutions aiming to promote sustainable behavior at the collective level.

5. Future directions

In this section, we present important open questions that can be addressed using neuroscientific approaches. We also introduce an emerging transdisciplinary framework that can potentially bridge laboratory experiments and practices in the real world.

5-1. What are neural bases of persistent behavioral changes?

For an institution to be effective in the real world, behavioral changes induced by the institution has to be persistent (i.e., long-lasting) (van der Linden, 2017). To examine long-term effects of experimental treatments on real-world sustainable behavior, we need to longitudinally collect real-world behavioral measures over certain periods of time (e.g., a few months). This can be done by occasional follow-up data collections, for instance via online experiments or smartphone apps monitoring daily sustainable behavior. An interesting question is what brain regions can predict long-term behavioral changes. If persistent behavioral changes induced by an experimental treatment are supported by implicit brain processes (e.g., emotional processing subserved by brain regions such as the amygdala), brain data could be a better predictor of long-term behavioral changes than self-report measures. It is also possible that some brain regions predict both immediate and long-term behavioral changes whereas other regions predict only immediate behavioral changes.

5-2. Are macroscopic changes of brain structure involved?

Another interesting neuroscientific question is what structural changes of the brain (i.e., brain plasticity) support behavioral changes related to intergenerational sustainability. If behavioral changes induced by a treatment are long-lasting, they should be accompanied by changes in the brain structure, either at a microscopic (e.g., synapses and spines) or macroscopic (e.g., cortical thickness of widespread areas) level. If the structural changes are macroscopic and large enough, they could be detected by structural MRI. Traditionally, intensive training of visuoperceptual or sensorimotor tasks has shown to induce macroscopic changes in task-relevant brain area (Zatorre et al., 2012). Recent studies have further revealed that interventional training aiming at enhancing cognitive ability (e.g., attention) and prosocial motivations (e.g., compassion) induces widespread changes in cortical thickness (Valk et al., 2017). In addition, cultural and environmental factors such as socioeconomic status (SES) modulate cortical thickness in widespread areas (even after controlling for genetic factors), which mediates effects of SES on cognitive ability such as executive functions (Noble et al., 2015). This raises the possibility that an immersive exposure to interventional training or educational programmes aiming at enhancing concern for future generations may induce widespread structural changes, particularly in regions implicated in social cognition and/or intertemporal decisions. Testing such possibility would provide useful insight into neural underpinnings underlying sustainable behaviour.

5-3. Can neuroscience offer better understandings of distinct prosocial motivations?

As we described before, prosocial motivations are multifaceted, and some of them may contribute to intergenerationally-sustainable decisions while others may not. Although we have argued that naturally-grown empathy might not be enough for achieving sustainability at the collective level, related (but possibly distinct) prosocial motivations such as compassion and loving-kindness may be enhanced with training (Lutz et al., 2008), which may promote concern for future generations by transcending perceived distances. In addition, impartiality (as opposed to parochialism) may play key roles in making intergenerationally-rational decisions (Baumgartner et al., 2013; Everett et al., 2018), which may counteract our natural biases toward the current generation.

Neuroimaging may allow us to clarify common and distinct substrates for these interrelated prosocial motivations, and help to understand neurocognitive components particularly important for intergenerationally-sustainable decisions.

5-4. Transdisciplinary research bridging laboratory experiments and real-world practices

Throughout this paper, we have emphasized the importance of developing and implementing social institutions to leverage concern for future generations that are effective in real-world situations. To achieve this challenging goal, we need to accumulate empirical evidence in both laboratory settings and real-world practices in a translatable manner. An emerging transdisciplinary framework, called "Future Design" (Saijo, 2015), aims at this goal by facilitating collaborations among researchers and citizens (including policy makers). For instance, Kamijo et al. (2017) showed that an institution called an "imaginary future generation" promotes intergenerationally-sustainable decisions in a laboratory setting (i.e., the ISDG). In an imaginary future generation treatment, some players in the current generation take the perspective of future generations, and discuss with other members in the current generation on behalf of future generations. Importantly, the essentially same institution has recently been used in practices in several local governments in Japan (Hara et al., 2019). This may offer a useful opportunity to examine effects of certain institutions in both well-controlled laboratory settings and real-world practices in actual policy-making processes. Future neuroscience research may take the advantage of such situations, for example by inviting the same participants and/or using the same institutions for both laboratory neuroimaging experiments and real-world practices for policy making. A key concept in Future Design is “futurability,” which is defined as the ability to derive happiness from deciding and acting to forego current benefits in order to enrich future generations (Saijo, 2015). Empirical research using neuroimaging may clarify the neurobiological underpinnings of this concept. For the conceptual uniqueness and recent progresses in Future Design, see Saijo (2020).

6. Conclusions

Intergenerational sustainability dilemmas lie at the heart of pressing issues in the contemporary society such as climate change. To solve these dilemmas, we need novel social systems to enhance the current generation’s concern for future generations, thereby achieving the transformation toward sustainable societies. Neuroscience may play unique roles in advancing the transdisciplinary research for intergenerational sustainability.

References

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, *15*(10), 670–682. https://doi.org/10.1038/nrn3800

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, *100*(401), 464. https://doi.org/10.2307/2234133

Aron, A. R. (2019). The Climate Crisis Needs Attention from Cognitive Scientists. *Trends in Cognitive Sciences*, *23*(11), 903–906. https://doi.org/10.1016/j.tics.2019.08.001

Aron, A. R., Ivry, R. B., Jeffery, K. J., Poldrack, R. A., Schmidt, R., Summerfield, C., & Urai, A. E. (2020). How Can Neuroscientists Respond to the Climate Emergency? *Neuron*, *106*(1), 17–20. https://doi.org/10.1016/j.neuron.2020.02.019

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally Interacting Minds. *Science*, *329*(5995), 1081–1085. https://doi.org/10.1126/science.1185718

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, *14*(11), 1468–1474. https://doi.org/10.1038/nn.2933

Baumgartner, T., Schiller, B., Hill, C., & Knoch, D. (2013). Impartiality in humans is predicted by brain structure of dorsomedial prefrontal cortex. *NeuroImage*, *81*, 317–324. https://doi.org/10.1016/j.neuroimage.2013.05.047

Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T. R., Kane, M. J., & Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(5), 1087–1092. https://doi.org/10.1073/pnas.1713532115

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science (New York, N.Y.)*, *324*(5931), 1160–1164. https://doi.org/10.1126/science.1169694

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915. https://doi.org/10.1038/nature04981

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., Feigl, A. B., Gaziano, T., Mowafi, M., Pandya, A., Prettner, K., Rosenberg, L., Seligman, B., Stein, A., & Weinstein, C. (2011). *The global economic burden of noncommunicable diseases*. World Economic Forum.

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science*, *352*(6291), 1323–1326. https://doi.org/10.1126/science.aaf0784

Bruneau, E. G., Cikara, M., & Saxe, R. (2017). Parochial Empathy Predicts Reduced Altruism and the Endorsement of Passive Harm. *Social Psychological and Personality Science*, *8*(8), 934–942. https://doi.org/10.1177/1948550617693064

Bulley, A., & Schacter, D. L. (2020). Deliberating trade-offs with the future. *Nature*

Human Behaviour, *4*(3), 238–247. https://doi.org/10.1038/s41562-020-0834-9

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chua, H. F., Ho, S. S., Jasinska, A. J., Polk, T. A., Welsh, R. C., Liberzon, I., & Strecher, V. J. (2011). Self-related neural response to tailored smoking-cessation messages predicts quitting. *Nature Neuroscience*, *14*(4), 426–427. https://doi.org/10.1038/nn.2761

Cikara, M., & Fiske, S. T. (2011). Bounded Empathy: Neural Responses to Outgroup Targets' (Mis)fortunes. *Journal of Cognitive Neuroscience*, *23*(12), 3791–3803. https://doi.org/10.1162/jocn_a_00069

Crockett, M. J., Braams, B. R., Clark, L., Tobler, P. N., Robbins, T. W., & Kalenscher, T. (2013). Restricting temptations: Neural mechanisms of precommitment. *Neuron*, *79*(2), 391–401. https://doi.org/10.1016/j.neuron.2013.05.028

Crutzen, P. (2002). Geology of mankind. *Nature*, *415*(3).

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279

Decety, J., Bartal, I. B.-A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686). https://doi.org/10.1098/rstb.2015.0077

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, *18*(5), 767–772. https://doi.org/10.1038/nn.3981

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216. https://doi.org/10.1016/j.jesp.2018.07.004

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From Neural Responses to Population Behavior: Neural Focus Group Predicts Population-Level Media Effects. *Psychological Science*, *23*(5), 439–445. https://doi.org/10.1177/0956797611434964

Falk, E. B., Berkman, E. T., Mann, T., Harrison, B., & Lieberman, M. D. (2010). Predicting Persuasion-Induced Behavior Change from the Brain. *Journal of Neuroscience*, *30*(25), 8421–8424. https://doi.org/10.1523/JNEUROSCI.0063-10.2010

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., & Weber, E. U. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience*, *13*(5), 538–539. https://doi.org/10.1038/nn.2516

Fischer, M.-E., Irlenbusch, B., & Sadrieh, A. (2004). An intergenerational common pool resource experiment. *Journal of Environmental Economics and Management*, *48*(2), 811–836. https://doi.org/10.1016/j.jeem.2003.12.002

Fiske, S. T. (2002). What We Know Now About Bias and Intergroup Conflict, the Problem of the Century. *Current Directions in Psychological Science*, *11*(4), 123–128. https://doi.org/10.1111/1467-8721.00183

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–1543. https://doi.org/10.1126/science.1191883

Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game. *Experimental Economics*, *16*(2), 155–169. https://doi.org/10.1007/s10683-012-9337-5

Frederick, S., Loewenstein, G., & O’donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459–473. https://doi.org/10.1098/rstb.2002.1218

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351–1354.

Hara, K., Yoshioka, R., Kuroda, M., Kurimoto, S., & Saijo, T. (2019). Reconciling intergenerational conflicts with imaginary future generations: Evidence from a participatory deliberation practice in a municipality in Japan. *Sustainability Science*, *14*(6), 1605–1619. https://doi.org/10.1007/s11625-019-00684-x

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using Imagination to Understand the Neural Basis of Episodic Memory. *Journal of Neuroscience*, *27*(52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007

Hauser, O. P., Rand, D. G., Peysakhovich, A., & Nowak, M. A. (2014). Cooperating with the future. *Nature*, *511*(7508), 220–223. https://doi.org/10.1038/nature13530

Hein, G., Engelmann, J. B., Vollberg, M. C., & Tobler, P. N. (2016). How learning shapes the empathic brain. *Proceedings of the National Academy of Sciences*, *113*(1), 80–85. https://doi.org/10.1073/pnas.1514539112

Hein, G., Morishima, Y., Leiberg, S., Sul, S., & Fehr, E. (2016). The brain's functional network architecture reveals human motives. *Science*, *351*(6277), 1074–1078. https://doi.org/10.1126/science.aac7992

Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O'Doherty, J. P., & Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience*, *20*(8), 1142–1149. https://doi.org/10.1038/nn.4602

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

IPCC. (2014). *Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland.

Izuma, K., Aoki, R., Shibata, K., & Nakahara, K. (2019). Neural signals in amygdala predict implicit prejudice toward an ethnic outgroup. *NeuroImage*, *189*, 341–352. https://doi.org/10.1016/j.neuroimage.2019.01.019

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625–1633. https://doi.org/10.1038/nn2007

Kahn, M. E., Mohaddes, K., Ng, R. N., Pesaran, M. H., Raissi, M., & Yang, J.-C. (2019). *Long-term macroeconomic effects of climate change: A cross-country analysis* (No. 0898–2937). National Bureau of Economic Research.

Kamijo, Y., Komiya, A., Mifune, N., & Saijo, T. (2017). Negotiating with the future: Incorporating imaginary future generations into negotiations. *Sustainability*

Science, *12*(3), 409–420. https://doi.org/10.1007/s11625-016-0419-8

Kjellstrom, T., Maître, N., Saget, C., Otto, M., & Karimova, T. (2019). *Working on a warmer planet: The effect of heat stress on productivity and decent work*. http://www.ilo.org/global/publications/books/WCMS_711919/lang--en/index.htm

Klimecki, O. M., Mayer, S. V., Jusyte, A., Scheeff, J., & Schönenberg, M. (2016). Empathy promotes altruistic behavior in economic interactions. *Scientific Reports*, *6*(1), 31961. https://doi.org/10.1038/srep31961

Koechlin, E. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Kristal, A. S., & Whillans, A. V. (2020). What we can learn from five naturalistic field experiments that failed to shift commuter behaviour. *Nature Human Behaviour*, *4*(2), 169–176. https://doi.org/10.1038/s41562-019-0795-z

Langenbach, B. P., Baumgartner, T., Cazzoli, D., Müri, R. M., & Knoch, D. (2019). Inhibition of the right dlPFC by theta burst stimulation does not alter sustainable decision-making. *Scientific Reports*, *9*(1), 1–8.

Lenton, T. M., Rockström, J., Gaffney, O., Rahmstorf, S., Richardson, K., Steffen, W., & Schellnhuber, H. J. (2019). Climate tipping points—Too risky to bet against. *Nature*, *575*(7784), 592–595. https://doi.org/10.1038/d41586-019-03595-0

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109. https://doi.org/10.1037/0022-3514.37.11.2098

Lutz, A., Brefczynski-Lewis, J., Johnstone, T., & Davidson, R. J. (2008). Regulation of the Neural Circuitry of Emotion by Compassion Meditation: Effects of Meditative Expertise. *PLOS ONE*, *3*(3), e1897. https://doi.org/10.1371/journal.pone.0001897

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835. https://doi.org/10.1073/pnas.211415698

McClure, S. M. (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science*, *306*(5695), 503–507. https://doi.org/10.1126/science.1100907

Navajas, J., Niella, T., Garbulsky, G., Bahrami, B., & Sigman, M. (2018). Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. *Nature Human Behaviour*, *2*(2), 126–132. https://doi.org/10.1038/s41562-017-0273-4

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Libiger, O., Schork, N. J., Murray, S. S., Casey, B. J., Chang, L., Ernst, T. M., Frazier, J. A., Gruen, J. R., Kennedy, D. N., Van Zijl, P., … Sowell, E. R. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, *18*(5), 773–778. https://doi.org/10.1038/nn.3983

O'Neill, D. W., Fanning, A. L., Lamb, W. F., & Steinberger, J. K. (2018). A good life for all within planetary boundaries. *Nature Sustainability*, *1*(2), 88–95. https://doi.org/10.1038/s41893-018-0021-4

Park, S. Q., Kahnt, T., Dogan, A., Strang, S., Fehr, E., & Tobler, P. N. (2017). A neural link between generosity and happiness. *Nature Communications*, *8*(1), 15964. https://doi.org/10.1038/ncomms15964

Peters, J., & Büchel, C. (2010). Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. *Neuron*, *66*(1), 138–148. https://doi.org/10.1016/j.neuron.2010.03.026

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation. *Journal of Cognitive Neuroscience*, *12*(5), 729–738. https://doi.org/10.1162/089892900562552

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405. https://doi.org/10.1016/s0896-6273(02)00755-9

Rilling, J. K., & Sanfey, A. G. (2011). The Neuroscience of Social Decision-Making. *Annual Review of Psychology*, *62*(1), 23–48. https://doi.org/10.1146/annurev.psych.121208.131647

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. A. (2009). A safe operating space for humanity. *Nature*, *461*(7263), 472–475. https://doi.org/10.1038/461472a

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562. https://doi.org/10.1038/nrn3776

Saijo, T. (2015). Future design: Concept for a ministry of the future. *Social Design Engineering Serie*.

Saijo, T. (2020). *Future Design: Bequeathing Sustainable Natural Environments and Sustainable Societies to Future Generations*. http://www.souken.kochi-tech.ac.jp/seido/wp/SDES-2020-5.html

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science (New York, N.Y.)*, *300*(5626), 1755–1758. https://doi.org/10.1126/science.1082976

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic future thinking: Mechanisms and functions. *Current Opinion in Behavioral Sciences*, *17*, 41–50. https://doi.org/10.1016/j.cobeha.2017.06.002

Scholz, C., Baek, E. C., O’Donnell, M. B., Kim, H. S., Cappella, J. N., & Falk, E. B. (2017). A neural model of valuation and information virality. *Proceedings of the National Academy of Sciences*, *114*(11), 2881–2886. https://doi.org/10.1073/pnas.1615259114

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42,* 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*(3), 617–627. https://doi.org/10.1093/brain/awn279

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Sherstyuk, K., Tarui, N., Ravago, M.-L. V., & Saijo, T. (2016). Intergenerational Games with Dynamic Externalities and Climate Change Experiments. *Journal of the Association of Environmental and Resource Economists*, *3*(2), 247–281. https://doi.org/10.1086/684162

Shuman, E. K. (2010). Global climate change and infectious diseases. *New England Journal of Medicine*, *362*(12), 1061–1063.

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., & Ludwig, C. (2015). The trajectory of the Anthropocene: The great acceleration. *The Anthropocene Review*, *2*(1), 81–98.

Steffen, W., Crutzen, P. J., & McNeill, J. R. (2007). The Anthropocene: Are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment*, *36*(8), 614–621.

Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., Summerhayes, C. P., Barnosky, A. D., Cornell, S. E., Crucifix, M., Donges, J. F., Fetzer, I., Lade, S. J., Scheffer, M., Winkelmann, R., & Schellnhuber, H. J. (2018). Trajectories of the Earth System in the Anthropocene. *Proceedings of the*

National Academy of Sciences, *115*(33), 8252–8259. https://doi.org/10.1073/pnas.1810141115

Stolier, R. M., & Freeman, J. B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature Neuroscience*, *19*(6), 795–797. https://doi.org/10.1038/nn.4296

Stott, P. (2016). How climate change affects extreme weather events. *Science*, *352*(6293), 1517–1518.

Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., & Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences*, *112*(5), 1619–1624. https://doi.org/10.1073/pnas.1414715112

Suddendorf, T., Bulley, A., & Miloyan, B. (2018). Prospection and natural selection. *Current Opinion in Behavioral Sciences*, *24*, 26–31. https://doi.org/10.1016/j.cobeha.2018.01.019

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow$^{\text{TM}}$: Using Behavioral Economics to Increase Employee Saving. *Journal of Political Economy*, *112*(S1), S164–S187. JSTOR. https://doi.org/10.1086/380085

Timilsina, R., Kotani, K., Nakagawa, Y., & Saijo, T. (2017). *Can deliberative democracy resolve intergenerational sustainability dilemma?* 35.

Tong, L. C., Acikalin, M. Y., Genevsky, A., Shiv, B., & Knutson, B. (2020). Brain activity forecasts video engagement in an internet attention market. *Proceedings of the National Academy of Sciences*, *117*(12), 6936–6941. https://doi.org/10.1073/pnas.1905178117

Valk, S. L., Bernhardt, B. C., Trautwein, F.-M., Böckler, A., Kanske, P., Guizard, N., Collins, D. L., & Singer, T. (2017). Structural plasticity of the social brain: Differential change after socio-affective and cognitive mental training. *Science Advances*, 12.

van der Linden, S. (2017). The nature of viral altruism and how to make it stick. *Nature Human Behaviour*, *1*(3), 0041. https://doi.org/10.1038/s41562-016-0041

Wackernagel, M., Hanscom, L., & Lin, D. (2017). Making the Sustainable Development Goals Consistent with Sustainability. *Frontiers in Energy Research*, *5*. https://doi.org/10.3389/fenrg.2017.00018

Wade-Benzoni, K. A., Sondak, H., & Galinsky, A. D. (2010). Leaving a Legacy: Intergenerational Allocations of Benefits and Burdens. *Business Ethics Quarterly*, *20*(1), 7–34.

Watanabe, T., Takezawa, M., Nakawake, Y., Kunimatsu, A., Yamasue, H., Nakamura, M., Miyashita, Y., & Masuda, N. (2014). Two distinct neural mechanisms underlying indirect reciprocity. *Proceedings of the National Academy of Sciences*, *111*(11), 3990–3995. https://doi.org/10.1073/pnas.1318570111

World Bank. (2020). *Global Economic Prospects, June 2020*. DOI: 10.1596/978-1-4648-1553-9.

Wu, X., Lu, Y., Zhou, S., Chen, L., & Xu, B. (2016). Impact of climate change on human infectious diseases: Empirical evidence and human adaptation. *Environment International*, *86*, 14–23. https://doi.org/10.1016/j.envint.2015.09.007

Xiang, T., Lohrenz, T., & Montague, P. R. (2013). Computational substrates of norms and their violations during social exchange. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(3), 1099–1108a.

https://doi.org/10.1523/JNEUROSCI.1642-12.2013

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience*, *15*(4), 528–536. https://doi.org/10.1038/nn.3045

Figure legends

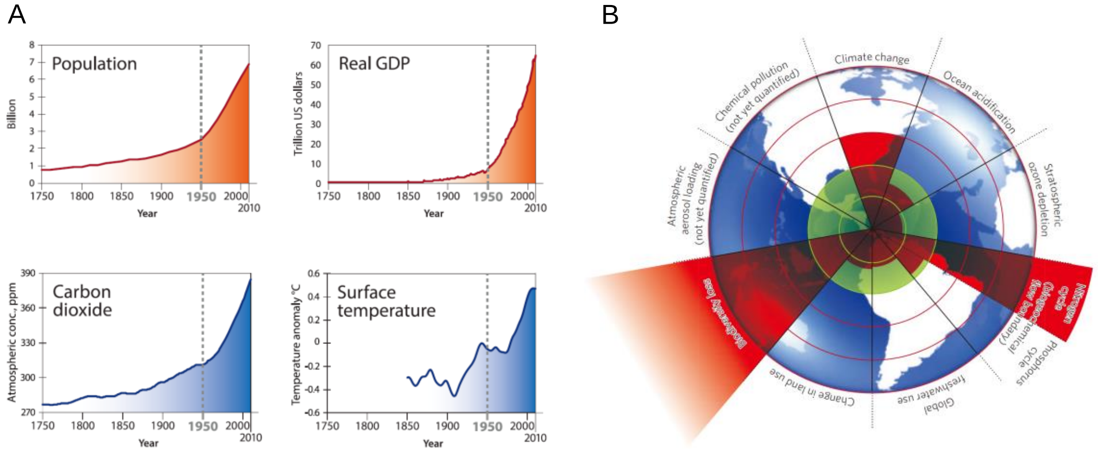


Figure 1. Scientific facts about climate change

A. Great acceleration. Increases in several socioeconomic indices (e.g., global population and real GDP) have rapidly accelerated after 1950 (upper panels). Changes in ecological indices (e.g., carbon dioxide in atmosphere and surface temperature) mirror this acceleration (lower panels). For a broader coverage of socioeconomic and ecological indices, see Steffen et al. (2015). Climate researchers generally agree with high confidence that the recent climate change is caused by human activity (IPCC, 2014). B. Planetary boundaries. Crossing certain biophysical thresholds may induce irreversible changes in the Earth's environment and endanger sustainability of humanity. The green area represents "safe operating space." The figure illustrates that human activity is approaching to or has already crossed the threshold in several domains of the Earth's ecosystem. The figures are adapted with permission from Steffen et al. (2015) and Rockström et al. (2009).

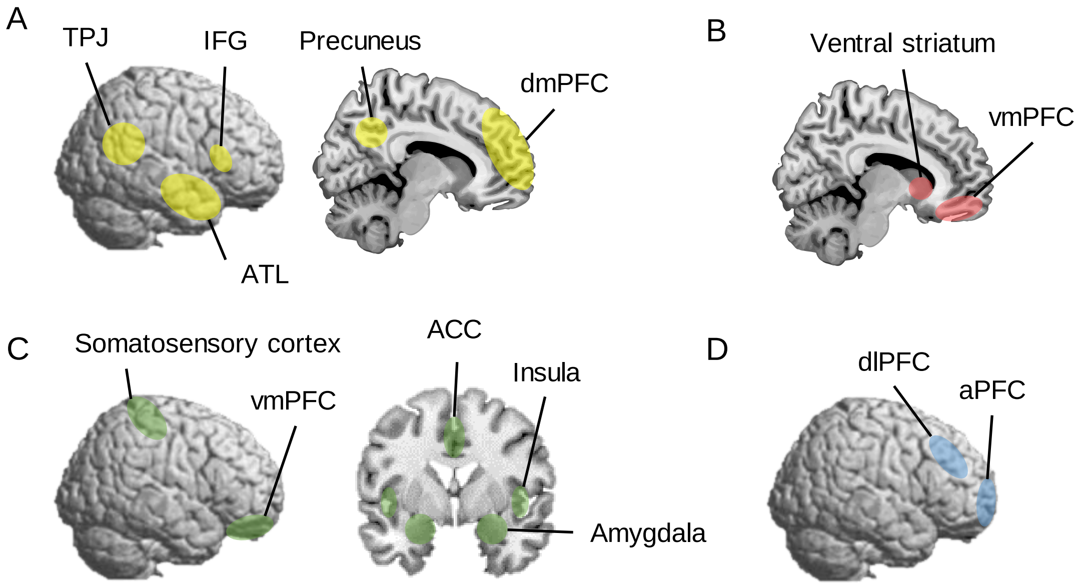
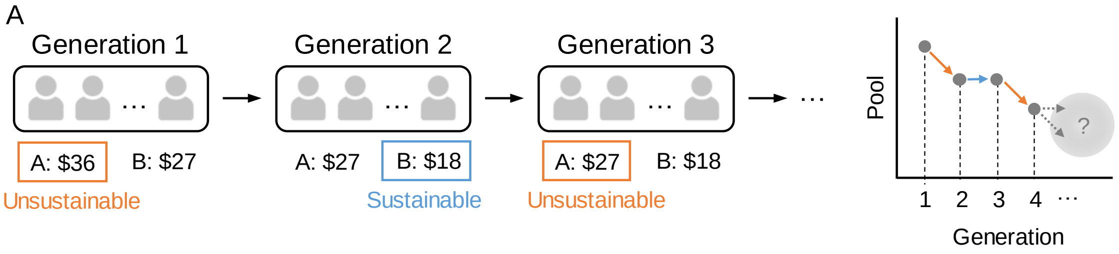


Figure 2. Possible neural substrates for intergenerationally-sustainable decision-making

A. Brain regions implicated in theory of mind (based on Schurz et al., 2014). B. Brain regions implicated in value-based decision-making and reward processing (based on Bartra et al., 2013). C. Brain regions implicated in affective empathy (based on Decety et al., 2016). D. Brain regions implicated in intertemporal decision-making and self-control (based on Crockett et al., 2013). Note that some regions appear in multiple panels (e.g., vmPFC), because a single brain region is often involved in multiple cognitive processes. TPJ: temporoparietal junction; IFG: inferior frontal gyrus; dmPFC: dorsomedial prefrontal cortex; vmPFC: ventromedial prefrontal cortex; ACC: anterior cingulate cortex; dlPFC: dorsolateral prefrontal cortex; aPFC: anterior prefrontal cortex.



B

Study	Prospective factor	Retrospective factor	Conversations
Hauser et al. (2014)	✓	-	-
Kamijo et al. (2017)	✓	✓	✓
Langenbach et al. (2019)	✓	-	-

Figure 3. Behavioral economic games

A. Illustration of a typical economic game (Kamijo et al., 2017) designed to study integenerationally-sustainable decision-making. A group of players represents a generation, and makes a collective decision that involves an intergenerational-sustainability dilemma (e.g., a trade-off between the current and future generations' benefits). The decision made by the current generation influences subsequent (i.e., future) generations but not previous (i.e., past) generations, reflecting the asymmetry of time in the real world. In this example, Generation 2 chooses the sustainable option, whereas Generation 1 and 3 choose the unsustainable option (which reduces the resources in the intergenerational common pool). B. Comparison among different games (Hauser et al., 2014; Kamijo et al., 2017; Langenbach et al., 2019). "Prospective factor" indicates whether players are informed about how their decisions affect future generations. "Retrospective factor" indicates whether players are informed about the history of decisions made by past generations.

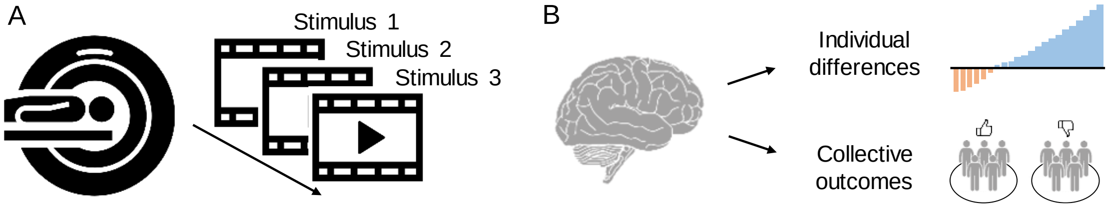


Figure 4. Predicting real-world outcomes from neural data

A. Participants are presented with naturalistic stimuli (e.g., movie clips that deliver messages promoting sustainable behavior or political announcements on sustainable policies) while their brain activity is measured. B. Patterns of neural responses (e.g., local activation, functional connectivity, and inter-subject correlation) could be related to either individual differences (e.g., behavioral changes induced by certain treatments) or collective outcomes (e.g., population-level responses to certain treatments). Short-term (immediate) and long-term (persistent) effects of a given treatment could be associated with differential patterns of neural responses.