Erasmus Journal for Philosophy and Economics,
Volume 17, Issue 1,
Summer 2024, pp. aa–aa.
https://doi.org/10.23941/ejpe.v17i1.849

The Challenge of Choosing Well: An Interview with Chrisoula Andreou

CHRISOULA ANDREOU (born in Montreal, Canada) is a Professor in the Philosophy Department at the University of Utah, an Executive Editor of the *Canadian Journal of Philosophy*, and an associate researcher at CRÉ (Centre de Recherche en Éthique) in Montreal. Her current research lies in the areas of practical reasoning, action theory, ethical theory, and applied ethics.

The *Erasmus Journal of Philosophy and Economics* interviewed Andreou on the occasion of the present special issue devoted to the philosophical debate on intra- and interpersonal dilemmas in ethics and rational choice, to which she has made prominent contributions. The interview covers Andreou's formative years (section I); her work on disorderly preferences, with a focus on her latest book *Choosing Well: the Good, the Bad, and the Trivial* (section II); the insights that can be gained by comparing intra- and interpersonal dilemmas in ethics and rational choice theory (section III); her view on the real-world applications of her philosophical work and the role of philosophers in society (section IV); her thoughts about the common thread that characterizes her past, present, and future research (section V); and, finally, her advice to young scholars (section VI).

I. Formative Years

EJPE: Professor Andreou, you did your BSc in Philosophy and Mathematics. What first drew you to these subjects? Why did the interest in philosophy eventually override the one in math? And how, if at all, has your training in math influenced your way of doing philosophy?

CHRISOULA ANDREOU: I took my first philosophy course while I was a

EJPE's NOTE: This interview was conducted by Annalisa Costella, Postdoctoral research fellow at the Vrije Universiteit (VU) Amsterdam and guest editor at *EJPE*, and Tessa Supèr, PhD student at the Vrije Universiteit (VU) Amsterdam, working on rationality in temporally extended and collective action.

Footnotes throughout the interview are editorial annotations directing readers to the relevant literature discussed in the text.

student in the Quebec CEGEP system[1] and was excited to discover that the kind of thinking I often found myself engaged in was a discipline that I could study. I enjoyed theoretical mathematics because of its connection with analysis and proofs. (I also assumed it was a more realistic basis for employment.) My undergraduate training in math allowed me to feel confident taking graduate level courses in game theory, to which I was also immediately drawn.

You also have a certificate in Hellenic Studies obtained during your years in Quebec's CEGEP system. Did your Greek origins foster your interest in the topic? And how, if at all, did this interest connect with your parallel interests in Math and Philosophy?

It was actually my first philosophy course, which was on ancient philosophy, that sparked my interest in Hellenic Studies. Though I grew up in a Greek community, very much steeped in 'everyday' Greek culture (food, dancing, language, traditions, etc.), philosophy and academic inquiries were not very visible to me. The Greek community I was most familiar with as a child was heavy on restaurant and factory workers keen on getting their children access to educational opportunities that they never had.

You did your PhD at the University of Pittsburgh, which is now most renowned for its excellence in philosophy of science. Was the strength of the department the same back then, and, if so, why did you decide to do a PhD there given that your PhD thesis was on (meta)ethics?

I didn't go to the University of Pittsburgh with a very clear idea of what I wanted to work on, but the value theorists there were inspiring, and I knew after I completed my coursework that I wanted to work on issues related to the rational authority of morality.

Talking about your PhD, your supervisor was David Gauthier and you were one of his last students. How was it to work with him?

Working with David Gauthier was a great experience for me, and it was an incredible privilege to be one of his last students before his retirement in 2001. He supplied just the right amount of challenging skeptical feedback and encouragement with respect to the works-in-progress that I provided him. And when, while I was still working on my PhD, I temporarily moved

[1] A CEGEP is a post-high school, pre-university educational center that is specific to Quebec's educational system.

back to Canada while my partner worked in Toronto to pay off his student loans, David would meet me in Toronto to talk shop whenever he and Joan had one of their regular visits to top up on stuff they missed from Toronto, including (if I remember correctly) some kind of raw cheese that was—at least at the time—not readily available in Pittsburgh.

Who else (mentors, junior or senior peers) played a major role in the early years of your career?

I had many mentors early on, including, to name just a few: Jim Ring (Dawson College), who broached the unthinkable topic of pursuing a degree in philosophy with something like "Well, I assume you'll be getting a degree in philosophy, but where?"; Jim Brown (University of Toronto), who made grad school seem like a realistic possibility and provided extremely helpful practical guidance; Tamara Horowitz (University of Pittsburgh), who helped me in my early grad school days before David Gauthier became my supervisor; and Leslie Francis (University of Utah) who helped me navigate the profession when it seemed like a foreign land. In terms of invaluable discussions regarding half-baked ideas, Elijah Millgram and I have been swapping drafts to pick apart over coffee since I arrived at the University of Utah over 20 years ago.

Let us now zoom in on your PhD Thesis 'The Moral Grip', which tackles the question 'Why be moral?'. In a recent interview, you stated that working on that area led you to your current research focus on rationality and irrationality.[2] Why were you interested in the topic of being moral in the first place? And how does the topic connect to your current work?

I was, in light of the significance of morality, very keen on exploring skeptical arguments against, and defenses in favor of, the rational authority of morality. Although I had mixed feelings about contractarianism, I felt the force of Gauthier's view that, to really salvage morality against the skeptical threats it faces, one must provide a theory of morality that has two features: it acknowledges that morality may require genuine sacrifices on the part of an agent; and it vindicates the idea that the demands of morality are, nonetheless, genuinely reason-giving. Without the first feature, morality is cast as more appealing and

[2] In August 2023, New Books Network interviewed Chrisoula Andreou about her latest book *Choosing Well: the Good, the Bad, and the Trivial.* The interview can be found at: https://newbooksnetwork.com/choosing-well.

accommodating than it really is; without the second, it emerges as involving a bluff. The focus of my dissertation was on critiquing the view that an agent's only reasons for action are provided by the agent's current 'subjective motivational set', which figures as a serious threat to the rational authority of morality. After my dissertation, I became increasingly focused on theories of practical reason and, eventually, on puzzles associated with rationality in intrapersonal cases. These intrapersonal puzzles can provide starker, more stripped-down variations of complexities found in interpersonal cases, particularly when it comes to the role of constraint in coordination and control.

You mention that, at the time, you had mixed feelings about contractarianism. Can you say something more about those feelings? And have you changed your mind since?
My mixed feelings were due largely to the fact that I found Gauthier's contractarianism, which was of particular interest to me in relation to debate about the rational authority of morality, counterintuitive in certain ways, and yet I was sympathetic to his core idea that constraint is important to rationality and, roughly put, to following rational courses of action or acting rationally over time. I still feel that way.

Let us close this first part of the interview with a question about the relationship between your research and yourself. In the preface of his book **Choosing for Changing Selves (2019)*****, Richard Pettigrew claims "A philosopher once said to me, while discussing the business of doing philosophy: 'We work on what we're bad at, on what we do poorly; we study the philosophy of whatever it is we find difficult.' [...] There are no doubt exceptions to this generalisation [...] but I'm not one of them" (Pettigrew 2019, 1). Similarly, in his famous paper 'Procrastination and Obedience', Nobel Laureate George Akerlof claims: "Some years back, when I was visiting India for a year, a good friend of mine, Joseph Stiglitz, visited me. [...] He left me with a box of clothes to be sent to him in the United States. [...] Each morning for over eight months I woke up and decided that the next morning would be the day to send the Stiglitz box. This occurred to me until a few months before my departure" (Akerlof 1991, 3). Do these quotes resonate with you? Was there something that you found puzzling about your own behavior that drove you to study the irrationality of temptation and procrastination?***

Well, my sense is that I'm no more of a procrastinator than the next person. My volume on procrastination, co-edited with Mark D. White (2010), was—ironically, I suppose—submitted on time. But I doubt I'm any less of a procrastinator than the next person either. The challenging choice and preference structures that give rise to temptation and prompt procrastination are everywhere. What varies, I suppose, is the domains in which we show control, the domains in which we let things slide, and the external constraints and supports we work within. Studying procrastination does, however, make it easier to spot. It's like when one first studies prisoner's dilemma situations and then starts noticing them everywhere. One of my usual forms of giving into temptation is wasting time seeking to maximize when I should be satisficing.

II. RATIONAL CHOICE AND DISORDERLY PREFERENCES

In your latest book,* Choosing Well: The Good, the Bad, and the Trivial *(2023b),[3] you focus on the theoretical and practical challenges for 'choosing well' which arise in the context of disorderly preferences (preferences that cannot be ranked from least to most preferred). For example, how to choose between sticking to a healthy diet and eating a delicious dessert that will only have a trivial effect on my health anyway? Or between a fun job with a low salary and a boring one with a high salary? Standard rational choice theory requires that an agent's preferences satisfy certain conditions, such as acyclicity and completeness. Your view differs: you hold that preferences can be rational without satisfying these conditions. Let's start with acyclicity. Traditionally, the requirement that rational agents cannot have cyclic preferences has been argued for by appealing to the money-pump argument.[4] However, you argue that there is an alternative interpretation of the money-pump argument which does not lead to the conclusion that cyclic preferences are inherently irrational. What do you consider the best argument(s) for either interpretation, and what convinced you to side with the latter?

What the money-pump argument shows is that, if an agent follows her cyclic preferences, she is susceptible to incurring a sure loss via a series of voluntary trades. It is often assumed, without argument, that if

[3] The money-pump argument shows that, if a person acts on her cyclic preferences, she could be repeatedly 'pumped' for money.
[4] See Davidson, McKinsey, and Suppes (1955).

following certain preferences is irrational, then having those preferences is irrational. But, although this assumption sounds plausible, it is hasty. If one's preference structure is appropriate (and that is, of course, a big if), all that is rationally required is that one exercise caution and not invariably act on one's preferences when this will be self-defeating. Putting aside cyclic preferences for the moment, we can see a clear and simple illustration of this requirement in EverBetter cases. In such cases, one is faced with an infinite series of ever better options, none of which one will get to enjoy if one does not at some point settle on an option (even though there will always be a better option that one is passing up in so settling). Given such a choice situation, it is not irrational to invariably prefer the next option in the series; it is, however, irrational to invariably follow one's preferences and never settle on an option. Since it is possible for a preference structure to be rational even if invariably acting on it would be self-defeating, the real question is: Can cyclic preferences ever figure as appropriate? In my view, there are cases in which the answer is 'yes'. The cases include ones in which, roughly speaking, individually trivial costs add up to an unacceptable cost. To illustrate in terms of the diet case, the 'cost' in relation to my appearance or my health of eating one more chip is trivial, and this is true of each chip I eat, but eating chip after chip whenever I am inclined to do so may have an unacceptable cost in relation to my appearance or my health. The result is appropriately cyclic preferences.

Given your view that (invariably) acting on cyclic preferences would be irrational, what is, according to you, the merit of insisting that having such preferences, on the other hand, is not irrational? Why not take a more neutral/agnostic stance with respect to the rationality of having cyclic preferences, if such preferences ought not to be acted on anyway? Is it simply because of the ubiquity of such preferences in our daily lives that we intuitively do not want to reject them as irrational, or is this commitment to rational/appropriate cyclic preferences in some way essential for (parts of) your theory?

The short answer is that it strikes me as dogmatic to dismiss cyclic preferences as irrational when they seem to be, not just familiar or common, but perfectly appropriate given the choice situation (as suggested in Quinn (1990)), and when there's no solid argument for the conclusion that they are irrational rather than just messy and difficult to navigate well. Ultimately, and as I say at the end of my book *Choosing*

Well, it's important that "rationality can handle quite a lot of messiness" since "rationality wouldn't be all that helpful if, whenever messiness threatened, we had to rush to its rescue rather than look to it for guidance" (Andreou 2023b, 175).

Next to acyclicity, the other main type of disorderly preferences you address in the book are incomplete preferences, which are taken to arise in cases of 'incommensurable alternatives'. You argue that even when options are neither better than one another, nor equally good, they are still comparable as "on a par" (Andreou 2023b, 73-90). You thereby reject the possibility of strictly incomparable options, and use this conclusion to defend the position that incomplete preferences can be rational. However, it seems that the rejection of strictly incomparable options could also lead to the conclusion that the existence of truly incomplete preferences should be rejected. Can you explain why your readers should not draw this conclusion instead?
I'm not committed to the view that cases in which two options are not rankable as one better than the other or as equally good are all cases of parity. Rather, my view, as I put it in the book, is that the distinction between relational and categorial subjective appraisal responses allows us to make sense of the possibility of two options being neither one better than the other nor exactly equally good, but still comparable as 'on a par'. Regarding incomparability, for complicated reasons that I won't get into here, I leave room for skepticism about incomparability without myself committing to such skepticism. In any case, I think that even if one rejects the idea that there are strictly incomparable options, the possibility of parity is enough to support the possibility of rationally incomplete preferences. This is because, according to the notion of parity that is relevant in the book, two options (for example, a slice of chocolate cake and a slice of berry pie) can be on a par for a certain agent given her choice situation even if they are not on a par for another agent given his choice situation; and when this is the case, it is appropriate for the former agent to have a preference gap in relation to the two options even if that is not appropriate for the latter agent.

In the same book, you argue that it is appropriate for a person to have incomplete preferences even if, in some cases, acting on incomplete preferences may make one vulnerable to being money-pumped. However, you continue, this vulnerability is not always problematic,

since a person may "proceed with care [and] avoid brute shuffling" (Andreou 2023b, 28). Can you expand on what your conception of 'proceeding with care' is? And how would this solution provide action guidance in the choice between options x and y, over which a person has incomplete preferences? Philosophers such as Ruth Chang (2013) seem to think that it is impermissible to pick. Rather, or so she seems to suggest, one is rationally required to choose. What do you think of this position?

When I say 'proceeding with care' I have in mind proceeding in a way that avoids self-defeating patterns of choice. When a person has incomplete preferences, proceeding with care involves not shuffling over time (via a series of choices between pairs of options that are not ranked by one's preferences) from one option to a worse option. Within the constraints of proceeding with care over time, picking between options that are not ranked by one's preferences is, I think, permissible, which is not to say that mourning the loss of the forgone good is out of place. I say more about this in my article 'Incommensurability and Hardness' (2024).

You address a paradox about self-defeating behavior: namely that behavior can only be truly self-defeating if the behavior is self-governed, but truly self-governed behavior can, inherently, never be self-defeating.[5] You solve this paradox by appealing to an agent's cyclic but stable values, which, if the agent always acts according to these values, can lead to truly self-defeating behavior, in the sense that it is also self-governed. You mention also the work of Michael Bratman, one of your main interlocutors on the topic of self-governance. Bratman (2009, 2018) agrees that, indeed, having conflicting values is unproblematic for self-governance. However, it is when an agent's intentions conflict that a breakdown in self-governance occurs. That is, a self-governing agent may value (as per your example) decent health as well as the pleasure of delicious desserts, but when she intends to act in line with both these values, there would be inconsistency between her intentions, which, according to Bratman, is a violation of one of the requirements for self-governance. For an agent to show the type of self-defeating behavior that you address, however, it seems that it must be true that she either (1) has indeed formed these inconsistent intentions (and acts on them) or (2) has managed to form consistent intentions, but fails to act accordingly (instead, she acts in line with certain values

[5] Andreou initially raised the paradox in Andreou (2012).

that she hasn't formed intentions to). According to Bratman, both would qualify as a breakdown of self-governance. So, following Bratman, one might have to conclude that self-defeating behavior that is also self-governed may be impossible after all. How would you respond to this worry? To what extent do you (dis)agree with Bratman's stance on self-governance?

Let me respond by appealing to the case of K in my book. K has an interest in gustatory pleasure and an interest in good health. Suppose that, given her interest in good health, K intends to avoid any gustatory indulgence that would have a significant negative impact on her health. The challenge is that, since no single potential indulgence has a significant negative impact on her health, K remains susceptible to self-defeating behavior. Now, one might insist that, to qualify as self-governing, K must have a more complicated intention: in particular, she must have a policy regarding culinary indulgences that is comprehensive in the sense that it settles, for each opportunity to indulge, whether or not to indulge. But it seems too strong to say that anything that one does that is not subsumed under such a comprehensive policy is not self-governed. I don't have a policy that determines how many chips I can eat per day, or per week, or per month, even though I really like chips but also aim to eat mostly healthy food. Is my chip-eating not self-governed? Must I have a policy that determines exactly how many chips (or even just how many empty calories) I can eat per unit of time for my behavior to qualify as self-governed? It might be responded that I only need a policy about whether I give my health or the enjoyment of chips more weight in my deliberations. But, as suggested above, this won't help in cases involving individually trivial effects. Like K, I can have a policy of avoiding any gustatory indulgence if it will significantly impact my health for the worse; but having one (more) chip (indeed, even one more bag of chips) will never have that power and so will never be precluded by my policy. To the extent that Bratman would allow that one can be self-governed even if one's only relevant policy is a vague policy (and my sense, based on some recent conversations, is that he is currently open to this idea), I think he may also be led to grant the possibility of self-defeating self-governance (at least relative to how I am using those terms).

III. THE PARALLEL BETWEEN INTRA- AND INTERPERSONAL DECISION-MAKING

Much of your work has focused on the interesting structural parallel that can be found between cases of intrapersonal and interpersonal choice dilemmas. In recent work (e.g., Andreou 2022b), you characterize some of the intrapersonal dilemmas that we all face (failing to save for retirement, not sticking to a dieting plan, and so on) as cases of intrapersonal free riding, and show that they are very similar to cases of interpersonal free riding, like the ones faced in 'tragedy of the commons'-type situations.[6] More generally, intra- and interpersonal choice problems may often be modeled as relevantly analogous. However, it does not seem necessarily true that the solutions to the problems need to thereby be analogous, especially if a proposed solution hinges on specific details which differ between intra- and interpersonal versions of the problem. To what extent do you believe the dissimilarities between intra- and interpersonal cases could affect the extension of potential solutions from one to the other?

I'm not sure that there are any essential dissimilarities between interpersonal and intrapersonal cases of free riding. And, as I've recently argued in my *Element on Commitment and Resoluteness in Rational Choice* (2022b), there is, in both cases, room for the relatively neglected possibility of 'impartially benevolent free riders', understood as free riders motivated by impartial benevolence. In such cases, the free rider seeks, out of impartial benevolence, to benefit the agent or self that she gets to decide whether or not to benefit—namely herself in the interpersonal case and her current self in the intrapersonal case—by sparing her from a burdensome contribution that will not make a significant difference in relation to a valued end.

Allow us to propose to you a potential worry regarding the analogy between intra- and interpersonal cases. You have shown that one way to potentially solve the theoretical and practical puzzles that arise in 'trivial effects cases' (where a sequence of acts with individually trivial effects together leads to a destructive outcome) would be to look differently at what is actually being done at each moment in which a trivial act takes place (Andreou 2014, 2022b). In short, you say that

[6] See Hardin (1968).

descriptions of what is being done at such moments are not limited to what fits in those moments, but can also include more temporally extended doings, which often do not have only trivial effects. You use this approach for both intrapersonal trivial effects cases, and cases that involve multiple agents. However, you mention that whether it is accurate to describe what is being done in terms of such a temporally extended description depends, in part, on whether an agent has certain dispositions. In the collective context, it must be the case that enough other agents also have similar dispositions. It might be objected that this creates problems for applying your approach to interpersonal cases: for example, what if agents are unaware of each other's dispositions? Or, what if not all individuals have the same concerns regarding the outcome they are collectively causing? Do you think that these differences between intrapersonal and interpersonal cases hinder the extension of your solution to the collective context? More broadly, do you believe there are other sources of disanalogy between intra- and interpersonal cases that may hamper the extension of your approach to interpersonal contexts?

Let me say first that I don't see any crucial difference here with respect to the task of establishing that a doing in progress can be harmful even if what is happening in the moment has no more than a trivial effect. In both individual and collective cases, a doing in progress can be harmful whether or not one is aware that it is going on. And both collectives and individuals may—sometimes excusably and sometimes inexcusably—fail to register what is going on because they fail to recognize, at least under the relevant description, that a certain disposition is in play. Moreover, like a collective, an individual can vary in terms of how unified or fragmented she is. If an individual is fragmented enough, she might, for certain purposes, be helpfully modeled as a series of time slices with conflicting concerns. In any case, my interest is primarily in free-rider cases (both intrapersonal and interpersonal) in which there is a dilemma despite the agents or selves in play sharing values. Interpersonal and intrapersonal/intertemporal cases of impartially benevolent free riding provide the starkest illustration of the possibility of such puzzling cases, though the cases can also include ones involving benevolent free riders who share values though their benevolence falls short of being completely impartial.

To what extent, if at all, do you think that it is a virtue of a solution to intrapersonal decision problems that it can be extended to interpersonal ones, and vice versa? If you agreed that some solutions to intrapersonal choice problems cannot be transferred to interpersonal ones, what do you think the implications would be? Would the 'impossibility of solution transfer' cast doubt on the solution itself or perhaps help shed light on the conceptual differences between a variety of intra- and interpersonal choice problems?
Insofar as there are any crucial disanalogies in relation to particular puzzles, a failure of solution transfer, or at least of complete transfer, would be expected and would not cast doubt on the solution. Of course, this is not to say anything about whether there are any such disanalogies. I'm not closed to this possibility, but I'm also not committed to it.

Several other prominent philosophers have also studied the similarity between intrapersonal intertemporal decision-making problems and collective action problems (such as Elster (1985), or Arntzenius and McCarthy (1997)). Is there any particular work that inspired you to study the analogy between such choice problems? And, if so, in what way does your work differ from theirs?
In terms of the analogy between intrapersonal decision-making problems and collective action problems, I think the work of George Ainslie (behavioral economist and research psychiatrist) played a big role in prompting me to develop an approach that complemented the prevailing approach. Like interpersonal dilemmas, intrapersonal dilemmas are commonly modeled as cases of conflict, as in Ainslie's work, with conflict existing between different individuals in the former cases, and between different time slices in the latter cases. My aim was to develop a complementary model that could accommodate puzzling cases in which the agents or selves involved could share the same values and preference rankings and yet still find themselves facing a dilemma. That is what is going on, I think, in many interpersonal and intrapersonal cases of procrastination wherein the choice situation prompts cyclic preferences.

In your most recent book, you chose to focus only on intrapersonal choice problems (rather than interpersonal ones too). Could you say something about your rationale for doing so? And to what extent do you think the novel analyses regarding self-defeating behavior and

disorderly preferences which you discuss in the book are also relevant in the interpersonal contexts where similar choice problems arise?
The rationale is largely practical. I do think that the analyses regarding self-defeating behavior and disorderly preferences that I discuss in the book are also relevant in interpersonal contexts, as is suggested in my work on environmental damage, in my previously mentioned *Element*, and in some of my work in progress, for example; but my aim in *Choosing Well* was to explore the significance of certain ranking structures and choice situations as straightforwardly as I could, and focusing on intrapersonal cases struck me as the way to go.

Your work seems to be driven by the belief that individuals are diachronically extended. To what extent, if at all, would your thinking about intrapersonal decision-making problems change if you were to be convinced by a time-slice—or fragmented (as you say in your previous answers)—conception of the individual?
I am open to the idea that modeling an individual as a series of time slices could be useful for certain purposes, but given that virtually all of our actions are temporally extended I can't really see how human action could be fully understood in terms of a time-slice conception of the individual.

IV. REAL-WORLD APPLICATIONS, POLICY-MAKING, AND THE ROLE OF PHILOSOPHERS

In your work, you often focus on societal problems and problems that we face in our everyday lives (such as Andreou (2006)). For example, you discuss strategies for how individuals could cope with procrastination in the context of starting to exercise more. To what extent do you think that your philosophical arguments can have a 'therapeutic' role by providing action guidance to individuals?
I think that improving our understanding of things like procrastination can be practically helpful because recognizing the structure of a problem can sometimes point to possible solutions. For example, knowing that procrastination is sometimes grounded in rationally permissible cyclic preferences, wherein there is no optimal option, can prompt us to be more accepting (other things equal) of measures that anchor our own behavior and the behavior of others to certain bright lines that, though artificially sharp, can help us to achieve acceptable patterns of choice over

time (measures such as developing fairly structured routines associated with eating, working, etc.).

You have hinted at the policy-making implications of your work on collective action problems in your article 'Environmental Preservation and Second-order Procrastination' (2007) where you bring up the idea of "laws requiring the formation of implementation intentions regarding the realization of environmental goals" (Andreou, 2007, 246), which could prevent second-order procrastination in environmental preservation. Do you see the potential for your philosophical work to inform policy-making decisions? Have you ever been involved in consultancy roles for policymakers or practitioners?
I think there is a lot of potential for philosophical work on decision making to inform policy-making decisions. For example, a few years ago, I was contacted by a creative solutions company working with the Social Security Administration on a project for improving retirement planning. My discussion with them focused on bright lines, anchor points, and the risks associated with complex choice situations that can prompt second-order procrastination. More recently, I have been invited to participate in an upcoming consultation for *The Lancet* Commission on 21st Century Global Threats to Health, which include (among other threats that policymakers and practitioners are grappling with) climate change, rising addiction, obesity, and worsening mental health.

Let us now zoom out from your specific engagement with societal problems to your broader opinion about the relationship between philosophers and society at large. What do you think the role of philosophers in society is? What kind of duties (if any) do they have towards the general public?
Well, it would be entertaining if I had a fun and dramatic blanket answer to this, like the Platonic answer that philosophers should wisely and benevolently rule and keep poets and other potential menaces in their place (ha ha). But I'll have to stick with mundane variations on the (still in some ways Platonic) idea that, like everyone else, philosophers should do their part in ways that fit with their skills and talents, and thereby hopefully benefit themselves and society. This can obviously look very different depending on their areas of inquiry and their extra-philosophical skills.

Talking about public engagement, as a scholar you fulfill both research and teaching duties. You likely are teaching or have taught on topics such as temptation and procrastination, which may be genuine challenges for many of your students. How, if at all, do you take the sensitive character of the topics you teach into account? How do students generally respond to these topics?
Not surprisingly, students tend to be quite interested in the topic of procrastination and, perhaps because it is so common and familiar, are usually curious and engaged rather than defensive when discussing models of procrastination and temptation.

More generally, what challenges do you think that a philosopher who engages with topics that are also of empirical importance is likely to encounter? For example, is the philosopher's work going to be somewhat dismissed by scientists working on the empirics of the phenomenon?
I hope not! The psychologists and economists that I have interacted with have generally not been at all dismissive, but my interactions are, of course, far from a random sample, since they have often been in the context of explicitly interdisciplinary workshops and projects. When it comes to interdisciplinary research, one substantial challenge is that of bridging different orientations and methodologies; but, given the value of being aware of and accountable to the insights of related research in other disciplines, the challenge is increasingly being taken up.

Are philosophers better equipped to engage with other disciplines than vice versa?
I don't know about that, but there may well be something to the idea that philosophy, given the inclusive nature of its areas of inquiry is, or should be, in some important sense, the most widely interdisciplinary discipline.

V. PAST AND FUTURE COMING TOGETHER

You have dedicated much of your initial work to modeling temptation as arising from cyclic preferences. What first made you realize that temptation problems could be modeled as such?
I think the realization was grounded in reflection on the role of individually trivial effects in certain cases of procrastination, combined with reflection on the connection between the accumulation of

individually trivial effects and cyclic preferences, all considered in light of Warren Quinn's (1990) intriguing puzzle of the self-torturer.

In the light of your description of temptation or weakness of will problems, you have proposed to solve these problems by invoking the two different kinds of appraisal responses: 'relational' and 'categorial'. You have argued that rational deliberation cannot lead a person to choose an option that is in a lower appraisal category than another option (Andreou 2015a). How did you first think about this solution?

I was led to this solution in part by thinking about why it might be that even folks who accept the possibility of rationally cyclic preferences think that, at least in some cases, ending up with certain options in a preference cycle is unacceptable. Since all options in a preference cycle are preferred to (at least) one option and dispreferred to another, this symmetry seems to speak against counting some options in a preference loop as permissible and others as impermissible. I was thus led to look for something that might break this apparent symmetry, even if our focus is exclusively on instrumental rationality and so on the agent's subjective appraisal responses. The possibility of our appraising options not just in terms of how they rank in relation to other options (relationally, that is to say) but also in terms of categories, such as, for example, 'great', 'fair', or 'bad', fit the bill. Insofar as relational and categorial subjective appraisal responses figure, in human psychology, as distinct types of responses, there is room for some options in a preference cycle to be appraised as in lower categories than other options (including when, as in the cases of particular interest in relation to many familiar cases of temptation, the number of categories in play is finite and the categories are neatly ordered); moreover, there is no reason to assume that instrumental rationality is accountable only to relational subjective appraisal responses, since categorial subjective appraisal responses are also the agent's own appraisal responses.

In what way is your solution distinct from, and to what extent does it complement the solutions proposed by, for example, Tenenbaum (2012, 2020) and Bratman (1987, 2018)?[7] And in what cases is it incompatible?

I think we're all in agreement that disorderly preferences can be, as Tenenbaum and Raffman (2012) put it, 'rationally innocent'. Tenenbaum

[7] EJPE refers to Tenenbaum & Raffman (2012), Tenenbaum (2020), Bratman (1987), and Bratman (2018).

seeks to solve the problem of how to proceed in such cases by appealing to the requirement that the agent opt for an alternative that is 'good enough' relative to her multiple different ends. My view, however—as I explain in my symposium paper (2023c) on Tenenbaum's 2020 book *Rational Powers in Action*—is that, tangential qualification aside, opting for an alternative that is 'good enough' can be irrational when there are also options that are, say, 'great'. Bratman's solution appeals to his 'no-regret condition'. But, like Holton (2009), I worry that Bratman's focus on regret is misplaced, in part because the question is not so much whether the agent will regret indulging but whether she should regret doing so, which brings us back to the question of how the agent should proceed. So the appeal to regret has not really helped us make much progress. Needless to say, these quick remarks can't possibly do justice to the complexity of the issues and of our points of divergence, which I say a great deal more about in my work. It also doesn't bring out the way in which our conversation is currently evolving in light of Tenebaum's and Bratman's contributions to the author-meets-critics APA session on my book *Choosing Well*.[8]

Your most recent writings include the topics of paternalism (Andreou 2018, 2023a) and incommensurable alternatives (Andreou, 2015b, 2020, 2021, 2022). What made you want to research those topics?
My interest in paternalism grew out of my reflections on strategies for self-control in the face of temptation, which led me to think about (the possibility of) self-paternalism. My interest in incommensurable alternatives fits right in with my overarching interest in choice situations that prompt disorderly preference structures in which the options cannot neatly be ranked from most preferred to least preferred, even allowing for ties.

Ruth Chang (2016) explicitly refers to your work on parity (Andreou 2012). Interestingly, she argues that, while you acknowledge that parity is an unconventional,* sui generis *value relation, she believes that your position is compatible with that of what she calls 'trichotomous incompatibilists', that is to say, philosophers who believe that cases of

[8] Andreou refers to the author-meets-critics Pacific Division American Philosophical Association symposium session on her book, *Choosing Well: The Good, the Bad, and the Trivial*, which was held in April 2023. Michael Bratman, Mark Budolfson, and Sergio Tenenbaum gave commentaries on her book.

'parity' are reducible to cases in which one of the three standard value relations (better than, worse than, or equal to) fails to hold. What do you think about her argument? And, in what way (if at all) does this misconstrue your position?

I am inclined to understand cases of parity as cases in which, roughly speaking, and putting aside important but tangential refinements, the options are in the same evaluative league or category (such as the category 'terrible') but are neither one better than the other nor exactly equally good. Chang, if I understand her correctly, thinks that, although this would make two options that are on a par comparable as 'in the same evaluative league or category', it would not suffice to qualify the relation 'on a par' as a pairwise comparative relation; parity would thus involve comparability (of a sort) but not necessarily pairwise comparability, and Chang is not keen on this because she is committed to the view that 'on a par' does qualify as a pairwise comparison. If this correctly captures Chang's concern about my view, then she is, I think, right that there is something that she is committed to establishing that I am not committed to. What I would emphasize is that, labeling issues aside, I am on board with the following views, which I think are both controversial and common ground between Chang and I: there are cases of comparability (in which the options are comparable as in the same evaluative category) that are not cases in which the options are one better than the other or exactly equally good; such cases are common; and such cases conflict with the orthodox assumption that, given any two options, they can be compared as one better than the other or as equally good. More recently, I have also made it clear that while some such options might qualify as "imprecisely equally good" (2022a, 71) there is also room for cases of the relevant sort that are not cases of imprecise equality. Moreover, I think my discussion of such cases can help shed light on Chang's view that there can be cases in which the difference in value between two options can have magnitude but not bias, which might seem hard to imagine.

Many philosophers who work on incommensurability and unconventional value relations have also contributed to debates in population ethics by analyzing the implications of their work for the infamous Repugnant Conclusion.[9] Given your interest in these topics,

[9] The Repugnant Conclusion, originally formulated by Parfit (1984), states that "for any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are

have you also engaged with these debates? And/or what other plans do you have for your future research?

I have broached the connection between my work and Parfit's infamous Repugnant Conclusion in my recent paper (Andreou 2020). There I argue that "a distinctive form of satisficing that involves reasoning in terms of leagues (understood as categories characterized by different grades or levels of quality) plays a crucial role in proper reasoning about what to do" (Andreou, 2020, 6), and can shed light on evading Parfit's Repugnant Conclusion.

You mentioned earlier that you have been recently pursuing another research strand on self-paternalism, a prima facie paradoxical notion. This is not the first time that you have argued for the existence of prima facie paradoxical concepts. As we discussed already, you also did it when you talked about self-defeating self-governance (Andreou 2012). Is there a relationship between the two topics?

Cases of self-defeating self-governance call for proceeding with care and showing restraint in relation to some of one's preferences, and self-paternalism can figure as an effective way of doing this, since, like paternalism more generally, self-paternalism can involve arranging things so that certain options are not as likely to be chosen as they otherwise would be.

Paternalism and self-paternalism are of interest to political philosophy and relevant to public policy. Do you plan to continue working on themes relevant to these latter areas of research inquiry?

Yes, I do expect to continue exploring the connection between my theoretical contributions regarding choice over time and topics that are broadly related to public policy and political philosophy. It's hard not to be drawn to this extremely interesting area of intersection.

If you were to describe your overall work, how would you say that it has developed over time? And what were the main drivers of your changing interests? What (if anything) is the common thread that has characterized your work so far?

My work is largely focused on rationality, irrationality, and preference structures and choice situations that raise challenges for effective choice

equal, would be better even though its members have lives that are barely worth living" (Parfit 1984, 917).

over time. My focus on rationality and irrationality has been there from the start, and the prominent place eventually accorded to choice over time is due to my increasing awareness of the often-neglected significance of the fact that our agency is exhibited in actions and courses of action that unfold over time. Further changing themes in my work are often generated by reflection on the significance of how we make choices and evaluate actions in relation to other philosophical topics.

VI. ADVICE TO YOUNG SCHOLARS

What advice would you give to PhD students who want to pursue a career in philosophy?
I guess if the focus is not just on engaging in philosophical inquiry but pursuing a career in philosophy, one may need to focus first on practical questions, such as: Will I need to go into debt and would doing so be feasible and make sense given my situation? Do I have personal commitments that would require things to work out promptly, neatly, and with a good amount of control over where I will live? If so, do I have a broad enough conception of what would count as a career in philosophy? If one has the luxury of engaging in philosophical inquiry, my advice would be to actively seek multiple mentors that will help one navigate challenges, to process failures as part and parcel of making progress, and to assume, by default, that every bit of feedback one gets on one's work can be helpful, even if only to illuminate where one might avert misunderstandings of one's position or approach.

What, in your opinion, are the ingredients for becoming a good philosopher and what are those to have success in academia, if the two differ?
Apart from a love for philosophical reflection and a great deal of curiosity, both seem to require a lot of patience, persistence, and resilience.

Do you believe that there is a trade-off between becoming an expert in one's own field of research as a PhD student and cultivating a broad variety of interests resulting in a somewhat dishomogeneous PhD thesis?
Well, one of the things I love about doing work on decision-making is that I feel like one can evade that bind as much as possible because of the way

in which the topic of decision-making intersects with a huge diversity of other topics. I often tell students who are doing coursework with me prior to their dissertation that, no matter what their interests, they shouldn't settle on a term paper topic that they can't get super excited about given just how much counts as fair game in relation to the topic of decision-making; and I invite them to brainstorm with me if they have doubts.

If you were given the opportunity to redo your PhD now, what would you change and what new challenges do you think you would face?

Well, it would be great if I was wiser then, but that's true now too, so I guess I'll just be grateful for having then enjoyed, and for still enjoying, the combination of internal and external factors that allowed me to press on with the aim of saying stuff that's new and sensible enough in conversation with others engaged in understanding the same philosophical issues. In terms of challenges I would face were I redoing my PhD now, it's hard to know how my studies would have been impacted by the way the profession and graduate studies in the field have evolved. If my answer above regarding the ingredients for being a good philosopher and being successful in academia is correct, and that's a big 'if', it may be that the challenges remain essentially the same.

REFERENCES

Akerlof, George A. 1991. "Procrastination and obedience." *The American economic review* 81 (2): 1–19.

Andreou, Chrisoula. 2006. "Environmental damage and the puzzle of the self-torturer." *Philosophy & Public Affairs* 34 (1): 95–108.

Andreou, Chrisoula. 2007. "Environmental preservation and second-order procrastination." *Philosophy & Public Affairs* 35 (3): 233–248.

Andreou, Chrisoula. 2010. "Coping with procrastination." In *The Thief of Time: Philosophical Essays on Procrastination*, edited by C. Andreou and M. D. White, 130–150. New York: Oxford University Press.

Andreou, C. 2012. "Self-defeating self-governance." *Philosophical Issues 22*, 20–34.

Andreou, C. 2015a. "The real puzzle of the self-torturer: Uncovering a new dimension of IR." *Canadian Journal of Philosophy*, *45*(5–6): 562–575.

Andreou, Chrisoula. 2015b. "Parity, comparability, and choice." *The Journal of Philosophy* 112 (1): 5–22.

Andreou, Chrisoula. 2018. "Self-paternalism." In *The Routledge Handbook of the Philosophy of Paternalism*, edited by Kalle Grill and Jason Hanna, 59–65. New York: Routledge.

Andreou, Chrisoula. 2020. "In a different league: Intransitivity, betterness, and league-based satisficing." In *Derek Parfit's Reasons and Persons*, edited by Andrea Sauchelli, 117–128. New York: Routledge.

Andreou, Chrisoula. 2021. "Incomparability and the huge-improvement arguments." *American Philosophical Quarterly* 58 (4): 307–318.
Andreou, Chrisoula. 2022a. "Parity Without Imprecise Equality." *Value Incommensurability*: 71.
Andreou, Chrisoula. 2022b. *Commitment and resoluteness in rational choice*. Cambridge: Cambridge University Press.
Andreou, Chrisoula. 2023a. "Paternalism and presumed superiority." *Analysis* 83 (1): 22–28.
Andreou, Chrisoula. 2023b. *Choosing Well: The Good, the Bad, and the Trivial*. New York: Oxford University Press.
Andreou, Chrisoula. 2023c. "Instrumental rationality and proceeding acceptably over time." *Philosophical Inquiries* 11 (1): 99–108.
Andreou, Chrisoula. 2024. "Incommensurability and Hardness." *Philosophical Studies*.
Andreou, Chrisoula, and Mark D. White, eds. 2010. *The thief of time: Philosophical essays on procrastination*. New York: Oxford University Press.
Arntzenius, Frank, and David McCarthy. 1997. "Self torture and group beneficence." Erkenntnis 47: 129–144.
Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Stanford: The Centre for the Study of Language and Information.
Bratman, Michael E. 2009. "Intention, practical rationality, and self-governance." *Ethics* 119 (3): 411–443.
Bratman, Michael E. 2018. *Planning, time, and self-governance: Essays in practical rationality*. New York: Oxford University Press.
Chang, Ruth. 2013. "Grounding practical normativity: going hybrid." *Philosophical Studies* 164: 163–187.
Chang, Ruth. 2016. "Parity: An Intuitive Case." *Ratio* 29 (4): 395–411.
Davidson, Donald, John Charles Chenoweth McKinsey, and Patrick Suppes. 1955. "Outlines of a formal theory of value, I." *Philosophy of science* 22 (2): 140–160.
Elster, Jon. 1985. "Weakness of will and the free-rider problem." *Economics & Philosophy* 1 (2): 231–265.
Hardin, Garret. 1968. "The tragedy of the commons". *Science* 162: 1243–128.
Holton, Richard. 2009. *Willing, wanting, waiting*. New York: OUP Oxford.
Parfit, Derek. 1984. *Reasons and persons*. New York: OUP Oxford.
Pettigrew, Richard. 2019. *Choosing for changing selves*. New York: Oxford University Press.
Tenenbaum, Sergio. 2020. *Rational Powers in Action: Instrumental rationality and extended agency*. New York: Oxford University Press.
Tenenbaum, Sergio, and Diana Raffman. 2012. "Vague projects and the puzzle of the self-torturer." *Ethics* 123 (1): 86–112.
Quinn, Warren. 1990. "The puzzle of the self-torturer." *Philosophical Studies* 59 (1): 79–90.