Review of Behavioral Economics, 2021, 8: 259–283

Inclusive Rationality: Struggle and Aspiration

Shruti Rajagopalan*

Mercatus Center, George Mason University, Fairfax, VA, USA; srajagopalan@mercatus.gmu.edu

ABSTRACT

In *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy* Rizzo and Whitman challenge behavioral economics and paternalism on multiple levels, from conceptual underpinnings and the meaning of rationality to more applied implications and policy recommendations. This paper delves deeper into Rizzo and Whitman's conception of inclusive rationality and places it within a larger historical tradition of purposeful behavior, internal conflict, and endogenous preference formation and change. I make a case to more carefully study two human characteristics – struggle and aspiration, to further the research program of inclusive rationality and behavioral economics. I argue that to extend this research program economists must: (1) have an idiom for struggle that does not deem a behavior rational or irrational either by assumption or by the normative standard set by an external expert; (2) take the process of endogenous preference formation and change more seriously; (3) have a language to model and explain aspiration or becoming; and (4) think about edge cases like relapsing addiction.

Keywords: Aspiration, inclusive rationality, struggle, behavioral economics, paternalism

JEL Codes: B29, B49, B53, D91

*I would like to thank Mario Rizzo and Glen Whitman for writing their excellent book. I also thank the organizers and the editors of the symposium for inviting me to contribute to this special conference and issue on the book. Finally, I thank Agnes Callard, Vishal Chaudhary, Nick Cowen, Tyler Cowen, Malte Dold, Nita Ghei, Boomer Kelley, Rohit Lamba, Prakhar Misra, and Shreyas Narla, and Biju Rao for helpful comments and suggestions.

ISSN 2326-6198; DOI 10.1561/105.00000142

Life is one long struggle in the dark. – Lucretius

1 Introduction

Rizzo and Whitman's *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy* challenges behavioral paternalism on multiple levels, from conceptual underpinnings and the meaning of rationality to more applied implications and policy recommendations. In this paper, I delve deeper into Rizzo and Whitman's [hereinafter RW] conception of inclusive rationality. I make a case to more carefully study two human characteristics – struggle and aspiration, to further the research on inclusive rationality and behavioral economics.

In their book, RW argue that the problem begins with behavioral economics and behavioral paternalism adopting the same restrictive definition of rationality that they are attempting to critique. In the process, behavioral economists neglect non-standard preferences, preference change, self-discovery, experimentation, and individuals adapting to their specific local context. To solve this conceptual error, the authors argue for a more inclusive definition of rationality, which they provide with great clarity and detail.

Most economists generally use rationality as a shorthand for technical definition or description of an individual agent's preferences and actions. When economists describe rational behavior, they mean an individual's preferences are reflexive, complete, and transitive. When economists adopted axiomatic rationality, their purposes were: first, to give preference theory and decision theory a logical foundation; second, to detail the necessary and sufficient conditions that must be placed on preferences to build a utility function; and third, to optimize that function given constraints, as part of theories with predictive or explanatory value.

RW (2019, 16) dub the neoclassical economics definition of rationality "puppet rationality" with little resemblance to human behavior. However, they are also sympathetic to its purpose. They argue that "these restrictive assumptions" of well behaved, transitive and complete preferences "were adopted not because they were especially plausible – from either a descriptive or normative perspective – but because they were analytically convenient. They made it easier to engage in mathematical modeling" (2019, 16).

Leading scholars in behavioral economics recognize this aspect of neoclassical economics. Thaler distinguishes between *Econs* – the abstract homo economicus individual used in rational choice axioms – with the more realistic *Human* (Thaler, 2015). He argues that economists get into trouble when they make highly specific predictions by assuming the world is inhabited by *Econs*.

Up to this point, there is little conflict between RW and the broader behavioral economics literature. Both critique the descriptive value of neoclassical models.

The divergence from RW emerges when behavioral economists, while criticizing and rejecting the descriptive accuracy of rational choice models, use rational choice theory as a normative standard, that is, the *right* or logically consistent optimizing model. In other words, though Thaler's behavioral economics inhabits the word of *Humans*, the normative standard of how *Humans should* act is to mimic *Econs*. As *Humans* we are misbehaving, against the standard "good behavior" of *Econs* (Rajagopalan, 2017).

Yet, neoclassical economic theory, no matter how reductive, was careful not to import any normative criteria for rationality. RW point out this fundamental divergence with the behavioral economists. In particular, the new paternalists implicitly or explicitly assume that the falsification of any particular claim of the positive model has normative import. They move from the descriptive or axiomatic to the prescriptive. To behaviorists, deviations from neoclassical models of rationality make individuals more human, but they also constitute errors in decision-making as measured against the standard of neoclassical economics or puppet rationality.

To correct this conceptual error, RW provide a broader and more inclusive definition of rationality that they call inclusive rationality.[1] In countering

[1]"Inclusive rationality means purposeful behavior based on subjective preferences and beliefs, in the presence of both environmental and cognitive constraints. This notion of rationality preserves the core notion of purposefulness, and in that sense it should seem familiar. But unlike other notions of rationality – many of which were invented for modeling purposes but have since taken on a life of their own – inclusive rationality does not dictate the normative structure of preferences and beliefs *a priori*. Instead, it allows a wide range of possibilities in terms of how real people select their goals, form and revise their beliefs, structure their decisions, and conceptualize the world. Their preferences and beliefs may be inchoate, incomplete, inconsistent, mutable, and dependent on context. Inclusive rationality can thus encompass choices and strategies that would not make sense under more restrictive notions of rationality. To be specific, real people may do all of the following and still qualify as inclusively rational:

- Experience internal conflict that has not yet been (and may never be) resolved;
- Have preferences that change over time;
- Have preferences that are indeterminate or incomplete – that is, that do not specify attitudes over all possible decisions at all possible times and states of the world;
- Have preferences that are in the process of being created or discovered;
- Have preferences that depend on context, including both the options available and the way in which decisions are framed;
- Hold beliefs that serve purposes other than truth-tracking, such as providing motivation or intrinsic satisfaction;
- Make inferences based not on the strict rules of classical logic, but on contextual and linguistic cues they have learned from human interaction;

both the neoclassical reduction of rationality as well as behavioral economists' use of neoclassical rationality as the normative standard, RW argue, that "[r]eal human beings, however, are not puppets. Their preferences and behavior may deviate from what is expected of agents in a model. But such deviation does not provide sufficient warrant for deeming them irrational" (2019, 16).

The major consequence of reliance on neoclassical rationality as a normative standard, and consequently recommending paternalist policies using that standard, is that this underestimates the capacity of individuals to solve their problems within their given social context and constraints. Simultaneously, this also overestimates the ability of policymakers to solve problems and design interventions because behavioral paternalists ignore knowledge and incentive problems that plague all policy, especially paternalistic policies. Paternalists are human, and subject to the same cognitive biases as those they are trying to correct, another important fact ignored by paternalists. RW argue that the policy recommendations by behavioral paternalists demand a level of knowledge that is impossible for policymakers to obtain given the usual constraints. And paternalists also assume away the incentives that plague the political process, especially interest group capture by special interests or other morally motivated paternalists.

In critiquing neoclassical economics for failing to accurately describe human behavior, behavioral economists miss the opportunity to actually provide a more descriptively accurate account of rationality. This is where RW are behavioral paternalists' most severe critics, but also provide the tools to be their most useful allies.

In this paper I delve deeper into the idea of inclusive rationality, the richness it adds to the behavioral economics literature, and its limits. There are two key aspects of human behavior – struggle and aspiration – that have received little attention in neoclassical or behavioral economics. In neoclassical economics, in equilibrium, the struggle is already over, or never occurred. In behavioral economics, the existence of internal conflict or human struggle is acknowledged,

---

- Indulge "biased" modes of decision-making when the costs are low and rein them in when the costs are high;
- Economize on scarce mental resources by refusing to impose perfect consistency on their preferences and beliefs;
- Structure their environments, possibly in ways that constrain their own choices;
- Adopt personal rules and resolutions that create internal incentive systems;
- Enlist the help of friends, family, and other groups to assist in attaining goals;
- Rely on institutions, social customs, and market structures to assist in attaining goals;
- Employ heuristics that minimize cognitive effort and/or informational input."

Rizzo and Whitman (2019, p. 27).

but it is in the background. This is because scholars treat any attempts or struggles, unless successful *ex post* and meeting the normative standard of neoclassical economics, as failings. Often, the reason is that inconsistency, especially one that violates the transitivity axiom, is immediately dubbed irrational. RW perhaps come closest, in their case for inclusive rationality, to understand internal conflicts and struggles.

Linked to struggle, but also distinct, is aspiration. Aspiring to be someone different necessarily involves a change in preferences (Callard, 2018). For instance, one may aspire to be someone who appreciates French noir cinema or run a marathon. One may not have those preferences yet and is hoping to form those preferences; this is a process that can involve some struggle. But endogenous preference change, while acknowledged, is not well incorporated into the theory of rationality in behavioral economics (Delmotte and Dold, 2021).

These specific human characteristics merit more research, especially by those who further inclusive rationality as a core research program within economics. As RW demonstrate, behaviors that may seem inconsistent at the first and maybe even second glance, might fit in a larger arc of rational behavior. Understanding struggle and aspiration better may also lead behavioral paternalists to reformulate the standards of what they consider inconsistent and irrational behavior.

Another question to consider is whether RW are too inclusive in their idea of rationality and their theory is non-falsifiable. Rizzo and Whitman (2020b) argue that their theory of inclusive rationality is falsifiable, in particular specific claims made within the larger research program. However, finding examples of falsifiability as external experts is difficult. Once an individual agent's motivations, knowledge, context, environment, etc. are all considered, can any action be called irrational?

Another reason to further examine struggle and aspiration more deeply is the fact that these aspects of human life show the difficulty in identifying irrational behavior when rationality is defined in a non-axiomatic sense. To struggle or to resolve that struggle because of some aspiration, is a process, and the process might be fraught with inconsistencies. To better understand the inclusive rationality program, it is important to identify and understand the difficult or edge cases between what is rational or irrational. These edge cases also point to the difficulty for external experts in finding irrational behaviors using the inclusive rationality paradigm. Studying edge cases might shed more light on this issue of falsifiable versus non-falsifiable theory and instead think in terms of individual falsifiable versus non-falsifiable claims. This is not because there are readily available examples of irrational behavior to falsify the theory of inclusive rationality. But because these edge cases point to the precise difficulties in finding irrationality, making inclusive rationality more important for development in future research.

2 Origins of Inclusive Rationality

Given that the origins of behavioral economics begin with looking for a more descriptively accurate theory and examination of human behavior, the first point of action is to look for a different, more realistic definition. Consider the summary RW give of the neoclassical notion of rationality:

> "Rationality can simply mean purposefulness, that is, trying to use the best means available to satisfy your goals or preferences given your beliefs about the world. It can mean taking an abstract approach to solving problems, applying universal systems of thought and inference, and following scientific methods. It can mean avoiding errors of logic and reasoning. It can mean revising one's beliefs in accordance with Bayes' Rule. It can mean having preferences that conform to certain axioms – transitivity, completeness, and so forth – which together guarantee the preferences are internally consistent and have a certain structure. In neoclassical economic theory, it has historically meant *all of the above*." (2019, 25, emphasis in original)

RW go back to an older and richer idea of rationality to mean purposeful behavior, as opposed to purposeful behavior that *only* fits an axiomatic description. Carl Menger, one of the cofounders of the marginal utility revolution, conceives of individuals within the market process as purposeful, but not infallible, and far from perfect. Jaffe (1976, 521) summarizes Menger's individual as "... a bumbling, erring, ill-informed creature, plagued with uncertainty, forever hovering between alluring hopes and haunting fears, and congenitally incapable of making finely calibrated decisions in pursuit of satisfactions."

Ludwig von Mises conception of rationality does not require some robot-like perfection. "It is a fact that human reason is not infallible and that man very often errs in selecting and applying the means. An action unsuited to the end sought falls short of expectation. It is contrary to purpose, but it is rational, i.e., the outcome of a reasonable – although faulty – deliberation and an attempt – although an ineffectual attempt – to attain a definite goal" (Mises, 1949, 44).

Smith distinguishes the neoclassical or constructivist rationality from ecological rationality, which deals with "adaptive human decision and with group processes of discovery in natural social environments" (Smith, 2008, 25). Simply put, it is the idea that context matters in individual choice. That context includes both interpersonal interaction and the social environment. In contrast, the neoclassical idea of rationality "purports to have universal validity irrespective of the individual's situation." (RW, 2019, 26).

Aside from the parallel neoclassical tradition having a fallible, though purposeful, human, other scholars in economics, psychology and political science have also incorporated a broader, more inclusive, definition of rationality as RW indicate. Gigerenzer (2008) asks psychologists and economists to replace individuals gifted with an omniscient mind, computing intricate probabilities and utilities, with individuals with a bounded mind – often reaching into an adaptive toolbox and relying on fast and frugal heuristics. With this replacement, one appreciates how the human mind adapts to situations, but also that in the face of incomplete knowledge and uncertainty, thinking fast leads to remarkably good results.

Gigerenzer *et al.* (1999) change the question from one of computational accuracy and clarity – where an objective answer exists for all individuals, to one of ecology. What is the subject's perception of the problem that he or she is trying to solve? With this approach, many of the biases, which are essentially behaviors deviating from the optimal model, disappear once the context becomes clear. Heuristics, instead of omniscience, become the appropriate, or rather the only tools, available under the circumstances.

Therefore, aside from purposeful behavior, rationality means successfully adapting to some task environment – that is, acting on the basis of beliefs, habits, and heuristics that function well in a specific environment to attain the agent's goals (Todd *et al.*, 2012). Economists have also incorporated the adaptive nature of the human mind, or how individuals learn by repeating interactions in a particular context, as demonstrated through experiments (List, 2004).

As Smith puts it, "when experimental results are contrary to standard concepts of rationality, assume not just that people are irrational, but that you may not have the right model of rational behavior. *Listen to what your subjects may be trying to tell you*" (Smith, 2005, 149, emphasis in original). And that is exactly what RW do throughout the book – they attempt to listen to the subjects that behavioral economists have deemed irrational. And then incorporate all these ideas and much more to their conception of rationality in the book that they call "inclusive rationality" and their definition and conception in its entirety (RW, 2019, 27). And typically, once they have moved away from the neoclassical understanding of rationality as the normative standard, much of the "irrationality" disappears.

There are some important characteristics distinguishing RW's approach to rationality from the other approaches in behavioral economics. They challenge the purpose and role of the representative agents as well as models used in neoclassical economics. First, if we only consider contributions towards improving the descriptive accuracy of neoclassical economics, few go farther than RW within economics. In this sense, they are perhaps the best allies to behavioral economics for further developing the tools and models to improve the descriptive accuracy and the predictive power of neoclassical economics.

But if the goal is to have a simplistic agent to abstract from reality to have pedagogical value, then the basic models of neoclassical models do the job quite nicely.

This is the reason I refrain from using the term "better" to describe RW's idea of rationality. It is not so much that one definition of rationality, axiomatic or inclusive, is better in some stand-alone sense. The question is one of analytical and pedagogical purposes. To what end is a scholar looking at the rationality of the individual? RW are acutely aware of this aspect of the definitions. Their criticism of behavioral economics is not that it imports neoclassical rationality, but that it shifts it from a positive and descriptive standard to a normative and prescriptive standard. They pertinently ask: "what is the significance of the *failure* of narrow rationality to predict or explain behavior in specific cases?" (2019, 38)

It is this shift of analytical goals that makes broader concepts of rationality, like ecological rationality, or RW's inclusive rationality important critiques of neoclassical economics as well as behavioral economics. It is not about a particular outcome, policy or paternalist intervention, but about a process by which economists, external to the agent, view the world. Smith argues that

> "if people in certain contexts make choices that contradict our formal theory of rationality, rather than conclude that they are irrational, some ask why, reexamine maintained hypotheses including all aspects of the experiments (procedures, payoffs, context, instructions, etc.) and inquire as to what new concepts and experimental designs can help us to better understand the behavior."

In other words,

> "[...] economists often label people "irrational" or "boundedly rational" without sufficiently emphasizing that this simply means that the individuals are not behaving according to one narrow stipulation of rational behavior. [...] economists are saying, in a roundabout fashion, that their theory of the phenomenon is wrong." (Smith, 2008, 40)

Finally, an important aspect of inclusive rationality is that it explicitly recognizes that preferences are not necessarily given and fixed. RW not only account for preferences that are not well-behaved, or do not fit standard technical definitions, but also account for endogenous preference change, preference creation, and incomplete preferences. This is an important point that is quite easily missed in technical models, even though individuals engage in preference formation and preference change on a daily basis (Delmotte and Dold, 2021).

3 Is Inclusive Rationality *too* Inclusive?

RW's conception of rationality includes a range of human characteristics and behaviors, irrespective of whether these characteristics and behaviors are recognized in formal models. However, one consequence of that inclusivity and breadth is that it raises the question – what kind of behavior, if any, would be considered irrational?

Inclusive rationality is thus far the richest and most comprehensive definition of rationality. It keeps the core of neoclassical rationality – purposeful behavior – while making it more descriptive of real-life human behavior as well as removing any normative import. RW genuinely try to understand the human condition, instead of conditioning humans to fit better into existing models. What, then, is the need to go out of our way to deem human behavior irrational?

In *The Logic of Scientific Discovery* (1959), Popper argued that the empirical truth cannot be known with absolute certainty. Even scientific laws cannot be verified beyond a shadow of a doubt: they can only be falsified by testing. One failed test is enough to falsify, but no number of conforming instances is sufficient to verify. Scientific laws are always hypothetical in character, and their validity remains open to falsification. Within economics, falsification, testing, checking and verification are the lynchpins of mainstream positivism (the method advocated by Friedman (1953)). Even economists like Bryan Caplan, otherwise sympathetic to the notion of inclusive rationality and its critique of new paternalism, argue that it is a theory that is "difficult or impossible to falsify" (Caplan, 2020).

The book does not provide examples of what is considered irrational under the conception of inclusive rationality. Therefore, on the one hand, neoclassical economists and behavioral economists are too quick to jump the gun and call all sorts of behaviors irrational, for the simple reason of deviating from *their* definition or model of rationality. But on the other end, is inclusive rationality, *too inclusive*? Is nothing irrational after considering limited cognitive capacity as well as an individual's specific context, institutions, social norms, as well as endogenous preference change and preference formation? If so, all choice, simply by virtue of it being the chosen action among the alternatives, is a rational choice. And viewed from this lens, inclusive rationality may be tautological and non-falsifiable.

RW are aware of this issue and anticipate the potential criticisms of non-falsifiability. They write, "positive claims should, in principle, be falsifiable. But some claims are more easily tested than others, and there is no guarantee that the most easily tested claims are also normatively relevant" (2019, 38).

While RW do not provide examples of specific irrational behaviors, they clarify that they do not believe that people are always fully rational and never engage in irrational behaviors. Their claim is not the absence of irrational behavior, but that it is very difficult to spot irrational behaviors once the

subjective and environmental context of the individual is fully considered (2019, 17).

This clarification is well taken. RW repeatedly detail the importance of the subjective preferences and specific environmental context of each individual while understanding or evaluating behavior and the difficulty this poses for the purpose of policy interventions. Because it is difficult to come up with examples of what would be considered irrational in the conception of RW, one important criticism is that the concept of inclusive rationality cannot be falsified like other positive claims of rationality.

They argue, and in fact encourage, scholars to test specific claims and examples of behavior that can be falsified, since it can help us better understand how people behave. However, they also caution, "that more restrictive positive claims are easier to test and potentially to falsify. For scientific purposes, then, more restrictive claims are often useful. But if the claims in question lack normative status, then although their falsification may tell us *more* about the world, it does not tell us that anything is *wrong* with the world. To make valid normative conclusions, we may need to consider less restrictive positive claims – which can be more difficult to test." (*Ibid*, 38–39)

Mises (1949, 66) argued that the theorems of economics are necessarily valid in instances where all the assumptions that were presupposed are given. But when those assumptions or conditions are not present the theorems have no practical significance. The question or criteria, then, is not falsifiability of a theory, but instead the *applicability* of a theory. Are the assumptions supposed in a given theory present such that they make the theory applicable? Answering this question is the job of the economist, who must understand the empirical evidence and determine the applicability of the theory.

If the purpose is to replace the definition used by neoclassical economists in their models, there is a problem of testability and falsifiability of RW's inclusive rationality. For this purpose, the authors, or others extending their work, must take on the hard task of coming up with specific examples of irrationality within inclusive rationality. While that might be one of the logical goals of a research program in inclusive rationality, that is not an area pursued by RW in this book.

If the purpose, as RW repeatedly indicate in the book, is to critique behavioral economics and in particular behavioral paternalism emerging from using neoclassical economics as a normative standard; then, their definition of rationality is not too inclusive. In fact, one may argue that in these instances, the more inclusive the idea of rationality, the better, because it will make the paternalists less likely to jump the gun and design interventions.

The second defense the authors offer is that the falsifiability claim might be taken too literally. The larger question at stake is not simply finding an example to falsify the theory, but whether the specific claims made under inclusive rationality are scientific and supported by evidence. And whether

inclusive rationality is applicable. Rizzo and Whitman (2020a), in response to Caplan (2020), argue that the way we think of scientific claims needs greater reflection. For many, "falsifiability is simply a shorthand way of asking *Is this claim scientific?* Or *Is this claim supported by evidence?*" (emphasis in original). They argue there "are mountains of evidence for inclusive rationality," *Escaping Paternalism* is in fact relentless in showing that people engage in all manner of self-regulatory behaviors. "Scientific hypotheses quite often generate predictions about the existence of certain phenomena. If a thorough search *fails* to turn up examples of such phenomena, then eventually that failure is deemed a falsification. . . . If the search for inclusively rational behaviors had turned up nothing, we would have had to reconsider our position. But it didn't turn out that way." (Rizzo and Whitman, 2020b) (emphasis in original).

This point requires further reflection even outside the context of the book. RW make a compelling argument that their conception of inclusive rationality is offering a research program, where individual claims may be confirmed or falsified, with a large amount of evidence overall confirming inclusive rationality. They explain the difficulties involved in proving examples of irrational behaviors with the limited tools as external economists/experts without knowing the true situation or context of the behavior. And that may well be correct, but this aspect also poses a problem.

Within the research program of inclusive rationality, in the book, the authors do not go far enough to explore the very difficult cases or edge cases. If the problem with finding examples of irrationality is the difficulty in categorizing a behavior as rational or irrational, then it is very important to give examples of those cases and showing why they might fall within the scope of inclusive rationality or considered irrational. Only by looking at these edge cases, can there be further development of inclusive rationality as a research program. I make such an attempt with two specific human characteristics – struggle and aspiration.

4 Struggle and Internal Conflict

When used in ordinary parlance, to struggle means "to have difficulty handling or coping with something" or "to make one's way with difficulty" or to "strive to achieve or attain something in the face of difficulty or resistance." Economists have not come up with a good definition for struggle. And this is not because struggling is not part of human behavior. They model the behavior before or after the struggle, rarely the process of struggling. Either it is encompassed in terms of higher costs, for example, in the meaning where it says to make one's way with difficulty. But thinking of problems in terms of costs represents struggle as a hurdle or a barrier that can be overcome at a higher cost, not internal struggle or conflict experienced by an individual that may never resolve.

Or it is assumed in the circumstances, usually considering uncertainty as a way of explaining the difficulty in handling or coping. Even scholars working within a decision-theoretic framework who recognize this kind of internal conflict caused by uncertainty over tastes and preferences model the problem as one where individuals are able to engage in costly contemplation before selecting an alternative (Ergin and Sarver, 2010).

One frequent example is of an individual struggling with the optimal diet, especially for those trying to lose weight to increase their fitness. Let us take the example of an individual, Susan, who very deliberately and intentionally wants to diet, but loves cookies. And in this particular instance, she chooses and eats the cookie when offered. The act of choosing the cookie is interpreted quite differently by economists depending on different views of rationality.

For a neoclassical economist studying consumer behavior, the act of choosing the cookie, reveals her preference for the cookie. However, the assumption of rationality is the foundation on which revealed preference axioms stand. So, in saying that she is rational in choosing a cookie over the diet, is assuming the problem away. Even if Susan ever had an internal struggle or conflict over that choice, it was fully resolved when the moment of choice arrived, and she chose the cookie.

If we push further and point to the fact that Susan said she wanted to diet and avoid cookies, neoclassical economists might simply dismiss it as cheap talk. Any talk or preference expressed that is not backed by choice, is just talk. But what is really going on is that the economist studies only the moment of choice, and in that moment, Susan chose a cookie, and therefore *must* prefer a cookie, *given that she is rational*. In the framework of revealed preference theory, preferences were simply descriptions of actual choices, without reference to mental states. That underlying mental preferences are revealed through choices. Here Susan's rationality is almost tautological.

In revealed preference theory, the economist assumes away the problem of struggle or internal conflict. This analysis does not differentiate divergent categories of individuals and their struggles – (1) those who choose cookies because they genuinely love cookies to eat it, no matter the calories (or perhaps because of the calories); (2) those who love cookies but would like to resist the temptation of eating the cookie; (3) those who are not sure whether or not they love eating cookies. Since in all three cases, the individual chooses the cookie, it is assumed that the individual preferred the cookie because axiomatic rationality is the underlying assumption. This is because the neoclassical economist is looking at the problem *after* the choice has been made, and the struggle, if any, was resolved just before the moment of choice through contemplation (Ergin and Sarver, 2010).

The behavioral economists, on the other hand, are quick to jump to the opposite conclusion. If Susan chose the cookie, after deliberately declaring the intention of wanting to stick to her diet, then she is deviating from the rational

and deliberate decision and therefore, acting irrationally. But first, behavioral economists, rightfully recognize that there is some kind of internal struggle or conflict. This is often described as the conflict between the two selves. Or as Kahneman phrased it, between thinking fast and slow (2011). The internal conflict, when resolved through a rational and deliberate thought process before the moment of choice, will lead Susan to resist the temptation of the cookie.

But it is not the recognition of the struggle that leads them to the opposite conclusion that Susan is irrational in choosing the cookie. The behavioral economists and paternalists automatically believe that thinking fast leads to worse outcomes than thinking slowly and deliberately, which is the normative standard. A slower, more deliberative process by Susan would have led to her resisting the temptation of the cookie. Many scholars have argued that myopic, nearsighted behavior can be attributed to "fast" choices and the use of heuristics (Rubinstein, 2003; Read *et al.*, 2013), and more farsighted behavior can be attributed to deliberative processing (Metcalfe and Mischel, 1999). Yet it is problematic to use thinking deliberately about the cookie as the normative standard to which Susan should aspire in order to be considered rational. This undermines the nature of the struggle, though in a different way than neoclassical rationality. In the case of the behavioralists, there is a struggle in choosing the better option because of temptation. But there is no question as to which of the two is the better option. Choosing the cookie while wanting to diet is an instance of myopia, or weakness of will, or inability to resist temptation.

The only way to understand if the behavior is in fact rational or irrational, is to know what Susan "really wants" irrespective of what she chooses. Philosopher, Callard (2018) points out that no one can really know which of the two is Susan's real or better preference, in order to judge if the choice deviated from it or not. Callard argues that akrasia,[2] that is, the weakness of will, is a form of intrinsic conflict and in order to qualify as akrasia, an individual must see the action as irrational, and she must perform it intentionally. Callard describes the internal dialogue of an akratic (Callard, 2018, 154–163) and points out that intrinsic conflicts have a perspective-dividing quality: seeing one option as valuable gets in the way of seeing the other as valuable. This is, again, because the self-monitoring activity connected to our valuing of one of the options requires us to reject the inclination to be attracted to the other option.

Callard deconstructs the internal dialogue of an akratic through this familiar example "I know I'll have more pleasure overall if I don't eat it . . . but the cookie is so tasty." (Callard, 2018, 162).

> "The "but" clause is not intended as a corrective to her earlier assessment of what she takes herself to know. It does not continue

[2]Callard argues that in order to qualify as akrasia, the action must possess both of these features: See Walker (1989) for a survey article on akrasia.

> or reopen her deliberations; rather, it pulls away from the whole project of deliberating. It represents a different thought from the one she had when she originally said to herself, "On the one hand, the cookie is tasty; on the other hand, I'm dieting." The difference in tone or emphasis marks a difference in the way the tastiness is apprehended, and this, in turn, corresponds to a difference in her reasons. The akratic's intrinsic conflict prevents these two reasons from being in conversation with one another. On the one hand, she deliberates in such a way as not to recognize the force of the fact that she wants pleasure NOW, which is to say, as though she were unproblematically committed to prudence. On the other hand, the point of view she adopts insofar as she experiences herself as having reason to indulge is, in turn, one from which she is immune to the force of the prudential considerations that strike her as having such deliberative significance." (*Ibid*)

For an external expert, there is no way to resolve the internal conflict conclusively favoring one alternative over another, without imposing the normative standard of the external expert. And this becomes clear as the behavioral paternalism literature tends to resolve internal conflict in favor of the expert's preferences and not the perspective of the individual in question (RW, 2019, 371). In contrast, inclusive rationality recognizes internal conflict and avoids the trap of picking one side over another in the internal dialogue. And the resolution of the internal conflict is also not a simple choice, but nested in a particular context where the individual has some self-awareness and anticipates the internal conflict.

Brennan and Buchanan offer the following example: "... consider Crusoe alone on his island (before Friday). He may deliberately choose to sleep on the beach at a location where the morning tide will rudely awaken him. By sleeping in such a place, Crusoe precommits himself to start the next day's work early. He closes off the option of deciding when to get up because his life plan includes work rather than sloth, and he wants to remove temptation of the latter." (Brennan and Buchanan, 1985, 72)

Rizzo and Whitman's book begins with Whitman's self-regulating efforts – starting with placing some snacks like pretzels and cookies on a higher shelf and escalating to banning ice-cream, that must necessarily live in a reachable freezer, altogether. That self-regulation that is required shows the nature of the unresolved conflict. The authors talk about how these kinds of strategies are both commonplace while also difficult to model with standard rationality axioms – like deliberately reducing the option set. Whitman's way of resolving the internal conflict when tempted with snacks is to avoid countering the problem by eliminating them from the choice set. The example talks about inclusive rationality and self-regulation.

Thankfully for Whitman, there was very little ongoing struggle (based on the description in the book) once this method of self-regulation was implemented. In fact, Whitman eliminated the ice cream, in anticipation of the struggle. But not all struggles can be eliminated by banishing ice cream from the reachable freezer. This is an important difference between these single instances of internal conflicts and ongoing struggles. Whether it is the literature in behavioral economics or on ecological/inclusive rationality, economists usually find ways to resolve internal conflict. Continuous, inconsistent, unresolved and ongoing conflicts and struggles are still under analyzed.

Let's say the problem is no longer a one-time internal conflict but an ongoing struggle, of say, portion control. The difficulty is no longer limited to certain snacks, like pretzels, cookies and ice cream but in exercising portion control altogether across most or all preferred foods in every meal or most meals. This is not an unusual dilemma for many. Obesity in the US among certain groups is often attributed to large portion sizes (Young and Nestle, 2002; Nestle, 2003). Banishing most foods altogether may not be a viable option. Consequently, some people who are on weight loss diets for months and years while also violating their diets for months and years. They manage to regulate some meals while dramatically failing at others. There are good days or weeks followed by binges and loss of control. While that might make sense for first time dieters, the odd thing about weight loss diets is that the longer the diet goes on the more likely they are violating it. This kind of process encompasses both struggle (of dieting) and aspiration (of becoming a healthier person).

Now those who start a diet might be on a path of self-regulation, which is part of inclusive rationality. Therefore, it is not irrational to start a weight loss diet. But my question, now, is about those who are always *trying*. Is it rational or irrational to be on a diet for months and years while simultaneously violating the diet for months and years, only prolonging the need for the said diet? After all, if they followed the diet plan, they may not have to diet for years. So, is it irrational to go on with diets when the individual is clearly failing at achieving the intended result?

As RW point out, the answer is not so simple. One needs to understand the particular circumstances and context of the individual. The length of the diet or inconsistent implementation by itself does not conclusively prove irrationality. But do people who try to lose weight for years and years act irrationally? They mean to lose weight, and they do succeed frequently by controlling their impulses. But they are also terribly inconsistent. Inconsistency in itself is not irrational in the framework of inclusive rationality (as shown by Whitman's method of self-regulation), but when does it become irrational? Is it irrational to eat potato chips while discussing the intention to lose weight in a chat with a friend?

One way is to think about rationality of the process from the lens of the outcome. If those who are inconsistently dieting, but still improving their

health outcomes or moving towards achieving their goals, then the process does not seem irrational. They may not follow the plan perfectly, but despite inconsistencies, it is helping them towards their final goal. But even if they are not improving their health outcomes with consistent dieting, the inconsistency alone might still not be sufficient to be irrational. Perhaps they would be in even worse health if they didn't even try the diet inconsistently. With weight loss, of course, whether an individual is getting closer to the goal, or at the very least not moving away from the goal, can be quite easily measured. But with other kinds of addictions or struggles, it may not be clear if during the process, that is, during the inconsistent and seemingly irrational behavior, the individual is moving towards achieving the outcome.

Let's take the case of addiction to substances like cigarettes or drugs. Two characteristics are essential for addictions: first, the more an individual partakes in an activity the more they want to partake in that activity; and second, the more an individual partakes in an activity, the lower their future utility. Though addictive behaviors are typically considered irrational, there are also models of rational addictions (Becker and Murphy, 1988). Becker and Murphy argue that individuals may engage in addictive choices but still make them, not because they are irrational, but because the gains from the activity exceed the costs of future addiction. This model has been applied to study addiction to cigarettes.

But there may be either individuals or substances that deviate from the rational addiction model used for substances. Maybe individuals initially believe it is a rational decision, but as the addiction progresses it may reduce their ability to control their consumption at a level where the benefits exceed future costs. While initially some may have used various tools of self-regulation to ensure that the benefits from consumption outweigh the costs, the nature of the addiction may be such that it weakens or eliminates their ability to manage their behavior.

Take the example of heroin addiction. Initially the consumption may have seemed rational and manageable, but the emergent addiction may progressively reduce the ability to engage in rational behavior to manage opioid consumption at a level where the benefits outweigh the costs. The origin point is rational, and some, including the individual struggling with the addiction, may argue that the end point is irrational.

With this acknowledgement, now there is a next stage where someone aspires to get treatment for their substance use disorder. But knowing the nature of the problem, it may not be easy. Is giving up the substance, when the addiction interferes with enjoying life, rational? Yes, that seems about right. But as Callard (2018) argues, a decision (like checking into rehab) is a step on the path to becoming someone else. But merely taking that decision will not complete their transformation to recovery: the path to success is not that simple.

But what is the appropriate process for giving up an addiction? Is relapsing rational or irrational? Is relapsing multiple times rational or irrational? What about those who relapse multiple times but also keep going back to therapy for their addiction. Is failing drug rehabilitation therapy and still going back for more rational or irrational? These are real struggles that thousands of people with substance use disorder experience every day.

Even though there is a high chance of relapse with substances like heroin, an individual who has confronted their substance use disorder and tried to give it up through rehab and therapy are engaging in rational behavior. They are ensuring that benefits from their current addiction does not lead to higher costs in the future, and therefore they are taking on costs in the present to enjoy benefits in the future by attempting to manage their substance use disorder. So, the goal to give up using the substance that is high cost in pecuniary and non-pecuniary terms may be a rational intention for an individual. In other words, the origin of this choice is rational. If the individual gives up the addiction successfully, then the decision to give it up was also clearly a rational choice when viewed from the end point or the point of the outcome.

One may, with deeper scrutiny, be able to gauge whether they are behaving rationally or irrationally at the start and end point of the struggle. However, for the process during the struggle, we don't have a good explanation for their behavior. At this point, the conception of inclusive rationality cautions against labeling an addict who is struggling and relapsing as irrational. Because the precise circumstances of the inconsistency or the internal conflict are not known to the outside expert.

That there is a struggle, does not mean it is irrational behavior. Because an important aspect of aspiration is that the direction in which an individual is headed in the moment, and that which is the end point, may be different than what they want at a given point. But what if someone struggles for years on end? What if they are in and out of rehab and therapy and keep relapsing into addiction? In the case of heroin use disorder, most outside experts as well as the individual experiencing the addiction and trying to give it up, may believe that learning to manage the use disorder is worth the struggle. So, the process is rational. Perhaps, a more ambiguous case can be more illustrative.

The years' long struggle to lose weight, or to become fit, or to give up an addiction do not have a clear line separating the rational from the irrational. The problem is not just one of more extensive investigation. The issue is that these cannot be abstract examples of a behavior. They depend on the individual, at a point in time, without the benefit of hindsight, often with enormous uncertainty, and a lot of inconsistency. And a large part of aspiration is the uncertainty associated with the *process* of aspiration.

These cases are not easy to deem irrational under inclusive rationality. But they are also not necessarily rational within the conception of inclusive

rationality. It depends. And more case studies and efforts to find the edge cases or those where rationality or irrationality reveals itself in a particular way, would be helpful. This research program exploring edge cases needs development. Inclusive rationality, as a framework, comes closest to recognizing that these struggles and aspiration are part of the human condition, but needs to be developed further to develop a theory of human struggle within the framework of rationality.

The decision theory framework, where preferences are exogenous and given and it is only a matter of choosing the best among the alternatives, is not a good framework when it comes to struggle. This kind of "choice" becomes less relevant to this process, except that individuals *choose to struggle*. Individuals choose to become someone different. And this links struggle to aspiration.

5 Aspiration and Endogenous Preference Change

Callard defines aspiration as "the distinctive form of agency directed at the acquisition of values." (Callard, 2018, 4) She explains that the "aspirant sees that she does not have the values that she would like to have, and therefore seeks to move herself toward a better valuational condition. She senses that there is more out there to value than she currently values, and she strives to come to see what she cannot yet get fully into view." (*Ibid.*, 5)

In the previous section I detailed the seemingly irrational person with substance use disorder who keeps relapsing but also keeps going back to rehab. Here it is not clear if the person with use disorder has a true preference to consume heroin, or the true preference is to rid himself of the addiction. If an individual keeps relapsing and consuming heroin, either due to internal conflict or the external, physical and chemical nature of addiction, the temptation is to explain the relapses in the language of preferences and constraints. The drug's chemical alters the mental process due to physical reliance, to such an extent, that there is no sensible way of evaluating the benefits in the present compared to the costs in the future. But in understanding the choices of a person with substance use disorder, Buchanan moves us away from the decision-theory framework. "How can the prevailing orthodoxy account for such a simple act as that by a person in quitting smoking .... Or going on a diet? If the person's utility function is unchanged, and if the constraints faced do not change, how can we account for such behavior" (Buchanan, 1979, 102–103).

Standard decision-theory framework is not well suited to deal with these questions. Buchanan argues that "choosing should be distinguished from behaving" because genuine choice involves "conscious selection from among alternatives" ([1969] 1979, 40). And that "behavior can be predetermined" but "choice, by its nature, cannot be predetermined and remain choice ... In a

wholly determinist universe, choice is purely illusory, as is discussion about choice' (*Ibid* 40, n. 3).

The fault is really with using decision theory to analyze human struggle and aspiration, as it considers preferences to be given and fixed and that they must be optimized within constraints. Buchanan is skeptical about the applicability of the standard model of man as a utility-maximiser to important realms of decision-making, precisely because it presupposes that people's conduct is a passive, deterministic response to their circumstances and thereby excludes creative choices. Similarly, Sen famously critiques homo economicus as a "rational fool" (1977).

Buchanan's individual has the ability to step back and evaluate and modify preferences to resemble more closely the kind of person they would like to be. Buchanan argues that behavior of this kind cannot be understood using standard choice theory. For the latter presupposes that people's preferences, and the utility functions that represent them, are stable, thereby ruling out the very changes that the artifactual man is bent upon making. Appadurai (2004, 67) argues that aspirations are "never simply individual" but are "always formed in interaction and in the thick of social life." Furthermore, not all individuals have equal capacity to aspire, and the capacity to aspire is itself rooted in a given socioeconomic and cultural context (*Ibid.*, 68).

Going back to the example of individual dealing with addition, Buchanan writes that the individual "can surely imagine himself or herself freed of the habit, with a transformed set of preferences that would not include any desire to smoke." If he acts to accomplish this goal by committing himself to a rule that prohibits smoking, then "he will find that he does become different from the person he was. His preferences shift; he becomes the non-smoker that he had imagined himself capable of becoming" (Buchanan, 1979, 103). Callard (2018) argues that while the existence of intrinsic conflict can be identified without reference to aspiration, the resolution of intrinsic conflict is linked to aspiration, or the process of becoming.

The human characteristic that could explain these seemingly inconsistent behaviors between relapse and rehab is aspiration. Aspiration is a resolution of the conflict of the akratic, and while the end point of the resolution is clear, each step on the way to that end point, or which direction an individual is headed at a given point, may be unclear. And therefore, using decision-theory and judging the rationality of the process using the end outcome is not illustrative. These seemingly bizarre and inconsistent behaviors can be explained if we view the individual like Frank Knight, as an "aspiring rather than a desiring being" (Knight, 1922, 472) and that "man is interested in changing himself, even to changing the ultimate core of his being" (Knight, [1942] 1982, 281). Buchanan's artifactual man (1979) is rooted within this Knightian conception of aspiring.

Like Buchanan and Sen who make a distinction between aspiration led preference change from the standard decision-theory framework used by most economists, Callard argues that "aspirants exhibit a distinct form of rationality that is not a matter of decision at all" (Callard, 2018, 27). Furthermore, that "large transformations in people's lives are rational though their rationality is not best captured through the framework of decision-making." (*Ibid*, 54)

Similarly, Callard talks about the *process* of becoming. She argues that there is a particular kind of rationality that individuals exercise when the available information about a situation is in the process of changing; or their desires are in the process of changing. Therefore, the rational aspect of what is taking place is not in the decision, because in some sense there is no single moment of decision. Instead, there is an extended process, characterized by learning, and a changing information set, a changing structure of desires, and becoming something else. Callard concludes that in order to see the rationality in the process taking place, the outside expert must see it over time, and not just an instant, as viewed by economists in decision-theory models.

There are two aspects to Callard's argument, and it is helpful to separate them. The first is that thinking of optimizing preferences that are exogenously given and fixed cannot conceptualize aspiration. Though endogenous preferences formation is now recognized more in economics, both standard neoclassical economics as well as behavioral economics have ignored endogenous processes of preference formation and more importantly endogenous preference change. The second is the temporal element or the process of this kind of change. Usually when there is a temporal element, economists enter into a comparative statics exercise where the two distinct points of equilibrium are compared with no heed to the process. Or if the utility from a given choice is uncertain, then models may use expected utility or intertemporal utility maximization. Aspiring to become someone else is far removed from this kind of reductive exercise.

On endogenous preference formation, Buchanan is quite radical (Dold, 2018; Lewis and Dold, 2020). For Buchanan, it is not the given preferences optimized under constraints that lead to the choice. Instead the moment of choice itself is a process of discovery, and the individual changes along with the constraints she faces and the choices she makes. "I am here advancing the more radical notion that not even individuals have well defined and well-articulated objectives that exist independently of choices themselves" (Buchanan, 1979, p. 111).

This eventually boils down to the distinction between optimization and genuine choice. Buchanan discussing process of choosing argues that, "Individuals do not act so as to maximize utilities described in independently existing functions. They confront genuine choices, and the sequence of decisions taken may be conceptualized, *ex post* (after the choices), in terms of "as if" functions that are maximized. But these "as if" functions are, themselves, generated in the choosing process, not separately from such process" (Buchanan, 1982, 5).

This idea of genuine choice is also central to Shackle's analysis:

> "The elemental thing we study is choice. If choice means anything, it means origination. The making of history (on however small a scale) is the making possible one path of affairs rather than another. By origination, I would say (and here take a decisive step outside all orthodoxy, even the Austrian) we ought to mean an act of thought as a first cause, so that choice in its essential nature is unpredictable in its effects, its sequel. Many "choices" are of course a response or obedience to habit or simple reckoning. By choice we ought to mean a momentous act of thought. If such an act is truly originative, it cannot be foreknown in character or timing, and thus we are essentially denied the power to specify the sequel of any present choice as a singular path." (Shackle, 1988, 206)

Buchanan argues that a person chooses, "from among many imagined futures, and he remains necessarily uncertain as to how that which he chooses will work out. He has a clear interest in seeing that the choice set, the set of alternative imagined futures, remains as open as is naturally possible, and, if constrained, that the constraints be also of his own choosing." (Buchanan, 1979, 111)

In their definition for inclusive rationality, RW explicitly include preferences that change over time; preferences that are indeterminate or incomplete – that is, that do not specify attitudes over all possible decisions at all possible times and states of the world; and preferences that are in the process of being created or discovered (2019, 27).

The process of becoming something else is also not in an institutional vacuum. For both Buchanan and Sen, liberty and freedom respectively, are central to this description of rationality and the broader human constitution. Sen highlights the reciprocal relationship between rationality and freedom, especially in the process of choosing among alternatives, or even seeking better alternatives, or becoming someone else. "Just as rationality is important in assessing freedom . . . freedom too is central to rationality" (Sen, 2004, 52). However, he does not interpret this rationality in its axiomatic form but in a broader fashion – "of subjecting one's choices – of actions as well as objectives, values and priorities – to reasoned scrutiny" (Sen, 2004, 4). Buchanan makes a similar call for liberty, "Man wants liberty to become the man he wants to become. He does so precisely because he does not know what man he will want to become in time. Let us remove once and for all the instrumental defense of liberty, the only one that can possibly be derived directly from orthodox economic analysis. Man does not want liberty in order to maximize his utility, or that of the society of which he is a part. He wants liberty to become the man he wants to become." (Buchanan, 1979, 112)

Following Buchanan, Knight, and Shackle's thinking on aspiration and endogenous preference formation and change diverges quite dramatically from

the standard decision-theoretic framework within economics. It's not even a matter of meta level decision making over preferences and constraints for decision-making. It is now a matter of genuine surprise and uncertainty over choice sets and entire paths or possibilities an individual might pursue. This individual will likely be inconsistent, relapse and recover alternately, and find a way to accomplish becoming a different individual. However, this individual, as Rizzo and Whitman argue, will also find a way to carve out a very individual and context specific path to accomplish his vision.

This framework must be extended to truly understand the behavior of a person with use disorder and their struggle with relapsing and rehabilitation. Or any similar edge cases, of behaviors that may nor may not easily be categorized as rational or irrational behavior. Each point of their journey may fall on the side of rational or irrational behavior depending on whether one uses axiomatic, behavioral, or inclusive rationality. But only a framework that takes serious genuine choice, preference formation and preference change, and aspiration, can truly hope to understand that behavior.

6 Conclusion

Rizzo and Whitman's *Escaping Paternalism: Rationality, Behavioral Economics, and Public* challenges the orthodox definition of rationality and argues that behavioral economics and paternalism fail because they adopt the same restrictive definition of rationality that they are attempting to critique. To remedy this failing of both neoclassical economics and behavioral economics and paternalism they provide a broad and rich definition of rationality – inclusive rationality. And in the process, they open up new possibilities for further research.

In this paper I have situated *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy* within the larger historical tradition on rationality and endogenous preference formation and change, and indicate directions to take the research on inclusive rationality further. Inclusive rationality is well poised to analyze the long-standing problems within behavioral paternalism – like addiction – while also understanding new and understudied areas in economics like struggle and aspiration.

To extend this research program economists must revisit four aspects of rationality. First, have an idiom for struggle that does not deem a behavior rational or irrational either by assumption or by the normative standard set by an external expert. Second, researchers must take the process of endogenous preference formation and change more seriously. Third, create a language to model and explain aspiration or becoming. And finally, study edge cases like addiction, which are difficult to model, but are also some of society's most

pressing problems. Inclusive rationality can and must be applied to these edge cases to better understand human behavior and action.

References

Appadurai, A. 2004. "The capacity to aspire: Culture and the terms of recognition". In: *Culture and Public Action*. Ed. by V. Rao and M. Walton. Stanford, CA: Stanford University Press. 59–84.

Becker, G. S. and K. M. Murphy. 1988. "A theory of rational addiction". *Journal of Political Economy*. 96(4): 675–700.

Brennan, G. and J. M. Buchanan. 1985. *The Reason of Rules: Constitutional Political Economy*. Cambridge: Cambridge University Press.

Buchanan, J. M. 1979. "Natural and artifactual man". In: *What Should Economists Do?* Ed. by J. M. Buchanan. Indianapolis: Liberty Fund. 93–112.

Buchanan, J. M. 1982. "Order defined in the process of its emergence". *Literature of Liberty*. 5(4): 5–18.

Buchanan, J. M. [1969] 1979. "Is economics the science of choice?" In: *What Should Economists Do?* Ed. by J. M. Buchanan. Indianapolis: Liberty Fund.

Callard, A. 2018. *Aspiration: The Agency of Becoming*. New York: Oxford University Press.

Caplan, B. 2020. "Escaping paternalism book club: Part 1". *The Library of Economics and Liberty*. URL: https://www.econlib.org/escaping-paternalism-book-club-part-1/.

Delmotte, C. and M. Dold. 2021. "Dynamic preferences and the behavioral case against sin taxes". *Constitutional Political Economy*: 1–20.

Dold, M. 2018. "Back to Buchanan? Explorations of welfare and subjectivism in behavioral economics". *Journal of Economic Methodology*. 25(2): 160–178.

Ergin, H. and T. Sarver. 2010. "A unique costly contemplation representation". *Econometrica*. 78(4): 1285–1339.

Friedman, M. 1953. *Essays in Positive Economics*. Chicago, IL: University of Chicago Press.

Gigerenzer, G. 2008. *Rationality for Mortals: How People Cope with Uncertainty*. New York: Oxford University Press.

Gigerenzer, G., P. M. Todd, and the ABC Research Group. 1999. *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Jaffe, W. 1976. "Menger, Jevons and Walras de-homogenized". *Economic Inquiry*. 14(4): 511–524.

Kahneman, D. 2011. *Thinking Fast and Thinking Slow*. New York: Farrar, Straus and Giroux.

Knight, F. H. 1922. "Ethics and the economic interpretation". *The Quarterly Journal of Economics*. 36(3): 454–481.

Knight, F. H. [1942] 1982. "Fact and value in social science". In: *Freedom and Reform: Essays in Economics and Social Philosophy*. Ed. by F. Knight. Indianapolis: Liberty Fund. 268–292.

Lewis, P. and M. Dold. 2020. "James Buchanan on the nature of choice: Ontology, artifactual man and the constitutional moment in political economy". *Cambridge Journal of Economics*. 44(5): 1159–1179.

List, J. A. 2004. "Neoclassical theory versus prospect theory: Evidence from the marketplace". *Econometrica*. 72(2): 615–625.

Metcalfe, J. and W. Mischel. 1999. "A hot/cool-system analysis of delay of gratification: Dynamics of willpower". *Psychological Review*. 106(1): 3–19.

Mises, L. 1949. *Human Action: A Treatise on Economics*. Indianapolis: Liberty Fund.

Nestle, M. 2003. "Increasing portion sizes in American diets: More calories, more obesity". *Journal of the Academy of Nutrition and Dietetics*. 103(1).

Rajagopalan, S. 2017. "Richard H. Thaler, misbehaving: The making of behavioral economics". *The Review of Austrian Economics*. 30(1): 137–141.

Read, D., S. Frederick, and M. Scholten. 2013. "Drift: An analysis of outcome framing in intertemporal choice". *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 39(2): 573–588.

Rizzo, M. J. and G. Whitman. 2019. *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy*. Cambridge University Press.

Rizzo, M. J. and G. Whitman. 2020a. "Escaping paternalism book club: Rizzo and Whitman response, part 2". *The Library of Economics and Liberty*. URL: https://www.econlib.org/escaping-paternalism-book-club-rizzo-and-whitman-response-part-2/.

Rizzo, M. J. and G. Whitman. 2020b. "Escaping paternalism book club: Rizzo and Whitman response, part 3". *The Library of Economics and Liberty*. URL: https://www.econlib.org/escaping-paternalism-book-club-rizzo-and-whitman-response-part-3/.

Rubinstein, A. 2003. "Economics and psychology? The case of hyperbolic discounting". *International Economic Review*. 44: 1207–1216.

Sen, A. 2004. *Rationality and Freedom*. Harvard University Press.

Sen, A. K. 1977. "Rational fools: A critique of the behavioral foundations of economic theory". *Philosophy & Public Affairs*: 317–344.

Shackle, G. L. S. 1988. *Business, Time, and Thought*. New York: New York University Press.

Smith, V. 2008. *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.

Thaler, R. H. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: Norton & Company.

Todd, P. M., G. Gigerenzer, and ABC Research Group. 2012. *Ecological Rationality: Intelligence in the World*. Oxford, UK: Oxford University Press/Cambridge, MA: MIT Press/Bradford Books.

Walker, A. F. 1989. "The problem of weakness of will". *Noûs*. 23(5): 653–676.

Young, L. R. and M. Nestle. 2002. "The contribution of expanding portion sizes to the US obesity epidemic". *Research and Practice*. 92(2).