

Taking Absurd Theories Seriously: Economics and the Case of Rational Addiction Theories*

Ole Rogeberg^{†‡}

Rational addiction theories illustrate how absurd choice theories in economics get taken seriously as possibly true explanations and tools for welfare analysis despite being poorly interpreted, empirically unfalsifiable, and based on wildly inaccurate assumptions selectively justified by ad-hoc stories. The lack of transparency introduced by poorly anchored mathematical models, the psychological persuasiveness of stories, and the way the profession neglects relevant issues are suggested as explanations for how what we perhaps should see as displays of technical skill and ingenuity are allowed to blur the lines between science and games.

1. Introduction. How can economists take absurd theories of choice seriously? And, why aren't these theories exposed to relevant criticism?

I discuss these questions through a case-study of rational addiction theory, currently “one of the standard tools in the economic analysis of the markets for drugs, alcohol, tobacco and other potentially addictive goods” (Ferguson 2000, 587). These theories are regularly featured in the leading economics journals, and their impact is such that the issues involved can not be brushed off as “mere” methodology by economists: Their standing in the literature means that they influence the thinking of economists on addiction, turn up in policy discussions such as Miron and Zwiebel 1995, filter out to non-economists through interdisciplinary

*Received August 2003; revised October 2003.

[†]To contact the author write to Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, N-0349 Oslo, Norway; e-mail: o.j.rogeberg@frisch.uio.no.

[‡]This work was financed by the Norwegian Research Council through a Ph. D. grant from HERO (Health Economics Research Programme at the University of Oslo). Thanks to two anonymous referees, my thesis supervisors (Professors Michael Hoel and Aanund Hylland at the University of Oslo) and Morten Nordberg for comments.

Philosophy of Science, 71 (July 2004) pp. 263–285. 0031-8248/2004/7103-0002\$10.00
Copyright 2004 by the Philosophy of Science Association. All rights reserved.

works such as Elster and Skog 1999 and Elster 1999a, and make economic theory seem irrelevant and absurd (see for instance Levine 2000).

Developed by Nobel Laureate Gary Becker and Kevin Murphy, the theory of rational addiction and its variants model well-informed and forward looking rational addicts with stable preferences, whose addictions are time-consistent, optimal consumption plans. The theory is claimed to *explain* the behavior of real addicts, i.e. identify the reasons why they do what they do.¹ It has also been used for *welfare analysis*.² I focus on two major rational addiction theories, but these features are common also to later variants such as Laibson 2001. These claims, taken together, commit these theorists to realism—the claim that their theories are true, that individuals are rational in the sense discussed and that the welfare effects observed in the theory are informative about the welfare effects that actually would follow for real individuals. I argue that this is an absurd claim.

The silliness of theories such as rational addiction theory raises the question of how they can be taken seriously. I suggest an explanation based on the relative impenetrability of mathematical arguments as compared to verbal arguments, the psychological feeling of understanding triggered by the stories included in articles, along with the neglect of some of the underlying issues by the economics profession.

To avoid misunderstanding: *There is nothing wrong with mathematical modeling per se*. Mathematics has proven itself an invaluable tool in a variety of sciences. The argument that follows simply states that mathematics—like any other tool—can be both applied and misapplied. The real problem with rational addiction theory is not that it uses mathematical modeling, but that it is empirically unfalsifiable, based on wildly

1. Examples from the two articles used in the following: Becker and Murphy 1988 states that “rational choice theory can explain a wide variety of addictive behavior” (676) and that “a theory of rational addiction does explain well-known features of addictions and appears to have a richer set of additional implications about addictive behavior than other approaches” (695). Orphanides and Zervos (1995) write that their theory substantially improves our “understanding of the determinants of addiction” (740). Of the general framework involved, Becker has stated that no “alternative approach—be it founded on “cultural,” “biological,” or “psychological” forces—comes close to providing comparable insights and explanatory power” into “a wide class of behavior, including habitual, social, and political behavior” (Becker 1996, 4).

2. Becker and Murphy (1988) state that “addictions, even strong ones, are usually rational in the sense of involving forward-looking maximization with stable preferences” (675) and that their model implies that addicts “would be even more unhappy if they were prevented from consuming the addictive goods” (691). Orphanides and Zervos (1995) use their model as the basis for a welfare analysis of information, education and public policy (751–754). See also the views on welfare analysis in Becker 1996, 20.

inaccurate assumptions, and poorly interpreted in a selective way that makes it seem plausible to some economists. Mathematics enters only by making this possible in practice.

Put differently, I am not saying economists should stop using mathematics, but that they should create a system of checks and balances by introducing further methods: Mathematics allowing for constraints and complex interdependencies where appropriate; a full, verbal explanation of the theory making it easier to think of contradictory evidence, weaknesses in assumptions and further factors that affect the mechanism or process discussed; statistical evidence used to establish the validity of empirical assertions. This is similar to other economist's desire for methodological pluralism and tailoring the method to the subject matter (e.g., Chick 1998).

Finally, I also discuss the possibility that the economists who author these theories do not *really* take them seriously in the way indicated by their written claims. Instead, they may be playing a theory-game that blurs the lines between economic science and game-playing, causing confusion and misunderstanding.

2. Use and Abuse of Mathematics in Economics.

Most of the topics on which economists hold views [that] are both different from 'common sense' and unambiguously closer to the truth than popular beliefs involve some form of adding-up constraint, indirect chain of causation, feedback effect, etc. Why can economists keep such things straight when even highly intelligent non-economists cannot? Because they have used mathematical models to help focus and form their intuition. (Krugman 1998, 1834)

Mathematical methods help us accurately formulate, reason, and examine certain types of arguments we are poor at handling unaided. This, as Paul Krugman notes, is a strong argument in favor of mathematics, though it also suggests caution when applying such techniques to individual decision-making. The mathematical model of decision making will be well-equipped to handle sophisticated, precise trade-offs caused by the many ways the value of goods might differ with our consumption of other goods past, present and in plans for the future. By Krugman's argument, even intelligent people would handle such trade offs poorly. We would be using a method that excels at X to model something that is poor at X.

A second implication of Krugman's argument is that our intuition and common sense will have a hard time following and critically assessing the explanation offered in mathematical models, precisely because these deal with deductions and implications that are hard to make in other ways. It becomes "possible, under the cover of a careful formalism, to make

statements which, if expressed in plain language, the mind would immediately repudiate” (Keynes quoted in Chick 1998, 1865).

In addition to making it hard for us to see the assumptions and implications of the mathematical model, human cognitive limitations also introduce a tendency to be overly swayed by stories. As recently discussed in Trout 2002, our subjective *feeling* of understanding drives our “judgments of the plausibility and, ultimately, the acceptability, of an explanation” despite being “in part the routine consequence of two well-documented biases in cognitive psychology: overconfidence and hindsight.” Roughly, if we are presented with evidence, it will often fail to have any influence in the absence of a good story, and if we are presented with a good story it will often have an excessive influence despite a lack of evidence (cf. Dawes 1999). I would argue that there is little beyond ad hoc stories in the way rational addiction theories are supported. Or, put differently, that stories can be told to support selected assumptions and results with such ease, that these theories all too easily create a mistaken impression of having accounted for a wide variety of facts.

Finally, the climate in the profession makes it hard to point out the absurdities made possible by the above factors. Economics seems to have embraced and identified itself with mathematical methods to the extent that economists in a recent “controversy” section of a top-ranking journal felt the need to argue that, while

[f]ormal techniques are powerful tools, [...] they can also be dangerous; the problem is to identify applications where they can be used safely. [...] My purpose is to make the case for such a debate to be opened. (Chick 1998, 1859)

The surprising thing is that even such (bland?) statements are not uncontroversial: In the same issue, Krugman denies flatly that “excessive formalism,” “excessive reliance on the deductive method,” and “arcane algebra that has no relationship to reality” is a problem in economics. The very best economists, he claims, as well as most of the successful ones, are free of such errors. Economists who think formalism is a problem are mistaken; outsiders who think so just dislike simple math that refute their pet doctrines (Krugman 1998).

While my remarks in support of methodological pluralism in the introduction may not differ much from the methodological guidelines Krugman ends up supporting in his piece, I believe—unlike him—that these are violated in practice by successful economists, with absurd theories discussed in full seriousness as the result. Rational addiction theory hopefully illustrates this.

3. Story-telling in Rational Addiction Theories.

3.1. Instrumentalism, Realism, and Idealization. Theories of rational addiction make assumptions concerning the choice rule, preferences and beliefs of people, and derive the resulting consumption plans. The choice rule is assumed to be rational, meaning that you use your beliefs to identify and choose the option best satisfying the preferences. Since there are no clear or obvious limits on preferences and beliefs, the set of possible rational consumption patterns remains to be identified. Currently derived patterns by rational addiction theories include rising, falling, shifting, cyclic, stable, stochastic, and chaotic, as well as endogenous quitting of both “cold turkey” and gradual kind.³ This flexibility is a strength in that the framework has a *prima facie* possibility of explaining these different consumption patterns. On the other hand, it makes it easy to seemingly explain something by just mimicking it. This case could be called ‘*mock explanations*’, in that such theories “explain” by reproducing observable regularities without capturing the causal mechanisms that produce these regularities, i.e. *describing* rather than *explaining*.

Some economists view such distinctions as unimportant, claiming support from the famous *as-if* methodology of Friedman 1953—usually interpreted as instrumentalism—which explicitly identifies prediction as the only aim of “positive economics.” Seen as instrumentalist, this would argue that good *as-if* theories are predictive devices, not explanations.

Granted, mock explanations may prove excellent predictive devices as long as the empirical regularities they describe remain stable, and the *as-if* stories help our understanding grasp and manipulate what would otherwise be disconnected facts and abstract relationships. As noted in Rosenberg 1992, 160, though, this defense comes at a cost: Since someone acting *as-if* he was rationally solving some decision problem would not behave optimally unless this was the actual decision problem he faced, assumptions matter when we turn to welfare analysis. Nor can *as-if* theories claim to explain in the sense of describing the mechanism or causal process underlying a phenomena, their aim is just to relate observable quantities in the simplest, most empirically successful way. To use a metaphor from Hausman 1992, an *as-if* theory of a car would be of no help to a mechanic if the car broke down. Nor, it would seem, could an *as-if* theory of addiction give insight into and understanding of addictions.

3. These examples are from Ferguson 2000 (rising, falling, stable, and shifting when new info), Feichtinger, Hommes et al. 1997 (chaotic), Becker and Murphy 1988 (rising, falling, quitting cold turkey, shifting, cyclic), Laibson 2001 (triggered by stochastic cues), Dockner and Feichtinger 1993 (cyclic), Adda and Lechene 2001 (rising, falling, endogenous quitting—all within a single model by changing parameter values).

Note that Friedman explicitly restricts himself to prediction and elsewhere explicitly warns against drawing inferences about welfare from a successfully predicting theory (Friedman and Savage 1952, 473). Since rational addiction theories are claimed to explain and provide a basis for welfare analysis, I infer that they are not just *as-if* theories.

While this disposes of the *as-if* defense, the question remains whether there are other ways to claim valid conclusions from invalid premises. Put differently: I assume in what follows that the validity of your conclusions derives from the validity of your reasoning and the soundness of your assumptions. Economists claim that their rational choice theories are a valid basis for welfare analysis and provide valid explanations of real people's behavior. I take this as a claim that their assumptions are "good enough" approximations to the truth. This has, indeed, been the defense of economic models invoked by some economists (see, for instance, Gibbard and Varian 1978 or Krugman 1998), and I attempt to show that this defense is invalid for rational addiction theories.

The discussion is related to idealization, "a literally false exaggeration that serves an abstractive or isolating theoretical purpose" (Hausman 1992, 132).⁴ Theories such as rational addiction theories could be seen, in such a light, as presenting idealized representations of real decision makers. Since we know that our idealized individual is a far cry from real individuals, this requires that we have a subsidiary theory concerning how deviations from this ideal will influence the outcome of our model. An example of such an attempt is the work of psychoeconomist Alan Nelson, who introduced a distinction between competence and performance along the lines of linguistic theory:

Discrepancies between the *competence* of economic agents based upon their ideal ability to maximize the utility function which we now assume they really do possess, and the actual *performance* of these agents in the marketplace (or even in the laboratory) can be accounted for in terms of interfering factors which are irrelevant to abstract microeconomic theory. It can be hypothesized that this interference will eventually be satisfactorily explained as resulting from things like misinformation, imperfect memory, limited calculating abilities, etc. (quoted in Rosenberg 1992, 137)

As Rosenberg (1992) notes in his discussion of Nelson, traditional economists show little if any interest in such work. This lack of interest is apparent in the rational addiction field, where a) direct validity is claimed

4. The argument in terms of idealization owes a large debt to one of the anonymous referees.

for explanation and welfare analysis, b) no discussion is included of where and in what ways the idealizations employed misrepresent reality, c) no interest is shown in and no method presented for examining the dependence of conclusions on the inaccuracies present. In other words: Yes, clearly untrue theories might have explanatory value and make true claims, but this needs to be established and shown, and that has not been done in the rational addiction literature.

It is of course conceivable that further arguments might be found that would make theories of the discussed kind valid for their proposed uses despite their absurdity, but good reasons for faith in this possibility are neither common knowledge nor given or referenced by articles of the kind under discussion.

3.2. Earlier Criticism of the Rational Addiction Approach. The main critic of rational addiction theory in the past has been philosopher Jon Elster, whose chief claim has been that Becker's theories are conceptually incoherent (see Elster 1997, 1999b, 2000b, 2000a). This claim has been contested by others (e.g. Rogeberg 2003). The sociologist Ole-Jørgen Skog (1999) has raised a number of good points concerning Becker and Murphy 1988 that were

partly to be read as a list of suggestions as to how their theory could be developed into a more elaborate theory [though also] some problems of a more fundamental nature, which are less easily resolved. (Skog 1999, 188)

Skog's more fundamental problems mainly concern the rationality assumed, which has also been criticized briefly in Winston 1980, 302. We return to this later.

While other kinds of criticism of the theory or earlier variants of the parent framework have been made (e.g., Cowen 1989, Rosenberg 1979, 1981, 1992), none of these have focused on bringing out the absurdity of the theories and the way mathematics is involved in hiding it.

3.3. Problems with the "Story-telling" Approach. A simple way to think of the mathematical model of rational addiction theory is as a formal system—a set of variables related to each other by a set of equations—that hopefully “mirrors” some real world system/mechanism/process. Our goal is a theory that captures some real world system/mechanism/process well enough that we can draw valid inferences from the theory (e.g., use rational addiction theory for welfare analysis) and explain the real phenomenon (e.g., understand addicts).

A theory can obviously not capture everything. For instance, economists often claim that mathematical choice models do not need to mirror the *process* by which humans identify the solution to their decision problem as long as they can specify the *outcome* that the human decision process will arrive at. If humans make optimal choices, any decision process able to identify this global optimum can be used by our theory (sometimes phrased by stating that procedural rationality is unimportant if people are substantively rational). While this argument suggests that our assumed solution method may be unimportant, it does require that *some* solution method equivalent to that used by our theoretical model is available to real people. More importantly, it requires that the decision problem itself is correctly specified. To say that people choose Y because it is *as-if* they were solving problem X is insufficient. To claim that X explains the choice of Y or establishes its optimality or welfare maximizing property, we must establish that people face a problem such as X.

To do this we need to '*anchor*' the assumptions in our formal system to their empirical counterparts.⁵ We interpret the model, state how it relates to the world, identify variables with empirical counterparts and give evidence that relationships specified between theoretical variables mirror relationships between the empirical phenomena. Unanchored assumptions are important to note, since a flexible framework may be made consistent with observations simply by altering these. This also brings out and catalogs inaccuracies introduced by our idealization of decision makers.

In practice, the anchors that tie the rational addiction theories to reality are stories. These range from incentives and mechanisms suggested by research findings to anecdotal evidence and common sense, and aim at establishing the assumptions as intuitively plausible and acceptable. The loosely structured way this is done in practice opens for three problems:

Weak anchors. Stories need to seem plausible and relevant so that they are rhetorically persuasive. Apart from this, there are no rules for "good" stories. They may suggest qualitative incentives without indicating either their strength or which groups of people we should expect to be aware of them. Instead of supporting some strange assumption X with a story, you may support some reasonable aspect of X or even some different but related assumption Y.

Unanchored assumptions. You do not need to provide explicit interpretations or tell stories to support the stranger assumptions of your theory. Weak anchors can make it seem as though they have been supported.

Unanchored facts. A story focuses our attention on an incentive or some aspect of reality. The set of possible stories that might be relevant is

5. I am not claiming this is *sufficient* to establish that the theory-explanation is correct, only that it is *necessary* for it to be a plausible alternative.

therefore immense, making it easy to tell a non-representative set of stories. If stories contradict assumptions, they can be left out. Stories that give the desired qualitative effects can be presented as though representative. Anomalies can be explained within the theory by telling further stories.

4. Problems in Practice and Their Importance.

4.1. Rational Addiction Theories. Drug users often experiment with drugs before gradually increasing consumption as they get hooked. Rational addiction theories use stable preferences to explain this unstable consumption by viewing present consumption partly as investment behavior. By smoking today you increase your smoking stock. Stable preferences are defined over such stocks in addition to the consumption goods.⁶

These extended preferences allow for various interactions. For example, a higher smoking stock is undesirable in itself (reduces health, produces withdrawal, etc.). It also makes some other goods more enjoyable, e.g. cigarettes (relief from withdrawal, learning to enjoy effects) and lunch with smoking colleagues (reduced sensitivity to tobacco smoke). Others may become less enjoyable, such as exercise (reduced cardiovascular fitness) and family parties (anti-smoking relatives).

According to Becker and Murphy (1988), consumption of heroin, tobacco, jogging, and television watching are all characterized by having such effects on future levels of various stocks. Individuals have precise, quantitative estimates both of these effects and of all future levels of “exogenous” factors such as prices, norms and public regulations. To decide whether or not to smoke tobacco, cigarettes, heroin, or crack, whether or not to turn on the television, go for a jog or eat a chocolate bar, individuals exploit these effects and design a detailed consumption plan for their future life exhausting all gains from all trade-offs across time and goods conditional on opportunities and exogenous factors; i.e. they map out—in precise detail—their optimal, lifetime plans of moment to moment consumption of these goods.

4.2. Detailed, Forward-looking, Long-term Planning. The defining assumption of rational addiction theories is that individuals make detailed,

6. The following is the simplest, most plausible interpretation of rational addiction theory. In Rogeberg 2003 I tried to interpret the entire extended utility framework, including the endogenous time preferences in Becker and Mulligan 1997 and Orphanides and Zervos 1998, but such interpretations seem less open to evidence and argue for agnosticism as regards results of welfare analysis.

forward-looking plans. This is the engine of the theories, needed for consumption to *rationaly* change over time despite stable preferences and no new information. It is also the feature that is used in Becker and Murphy 1988, 691–692 to distinguish this theory from other theories of addiction: “[I]n our model, both present and future behavior are part of a consistent, maximizing plan.”

In terms of Kukla 1991, this requires a maximal criterion of rationality, a logical omniscience where we immediately see all implications of our knowledge, an assumption that people always reason in the manner Krugman denied we are able to do without the use of mathematics. To find the solution, the economists turn to mathematical techniques such as the theory of optimal control and derive present consumption as the first step in a well-designed, coherent plan.⁷ Since you require your lifetime plan to know how much to consume today, changes in beliefs about future prices have immediate effects on present consumption. “A proposed hike in the tobacco tax may kick in next year, so I’m smoking a cigarette less daily to adjust my stock path.” This effect has been a major focus of the related econometric literature.

It’s a common saying that extraordinary claims require extraordinary evidence, so we might expect this central assumption to be well supported. Fehr and Zych (1998) found that people are poor at finding the correct solution to such decision problems in experimental settings. Skog (1999) seems to view the choice rule as descriptively inaccurate, but an ideal towards which we might strive. Others view the assumption of detailed planning of lifetime consumption paths as so obviously false there is no need to disprove it. An oft-quoted example is Winston 1980, 302. Since Winston’s article is cited for other reasons by Becker and Murphy (1988), they presumably knew this decision rule was seen as implausible. How do they justify it?

The answer is that they don’t. They state that their “paper relies on a weak concept of rationality” since individuals may strongly discount future utility (Becker and Murphy 1988, 683). This is a weak anchor, since it does not relate to the assumption of intricate, long-range planning that their rationality assumption represents. They also note that smokers do not neglect the future, since smoking was reduced when health dangers became known (687). This too, is a weak anchor, since there is a vast ocean between complete myopia and logical omniscience. This strange “unless you are oblivious of the future, you will be making intricate, long range plans” thinking is not a unique feature of Becker and Murphy, but also appears in the review of “The Economics of Smoking,” whose

7. See Ferguson 2000 for a clear exposition of the math involved.

authors support the assumption of forward-looking rationality by stating that

myopic behavior implies an infinite discounting of the future, while rational behavior implies that future implications are considered, while not ruling out a relatively high discount rate. (Chaloupka and Warner 2000, 1559)

There is an interesting neglect of *negative* evidence here: The assumption *should* be easy to support *if correct*. Since rational addicts have made forward looking plans, they are aware when they first take up smoking that “I’m smoking two a day now, next week I’m gonna smoke four, in a month I’ll be up to ten a day, and from October next year I’ll be smoking fifty a day till I quit at age 34.”

A second piece of negative evidence concerns the problems solved and presented in economic journals which seem, despite their sophistication, very simple compared to the full complexity a real person’s decision problem would have. This is strange: If real people routinely solve the full, complex problems, would we not have some way of communicating the reasoning behind our actions? Why would we find descriptions of the decision problem and its solution so bizarre? Why would simplified versions seem complex? To point to a market “anomaly”: If people routinely solve and implement the most fantastically sophisticated plans derived from accurate knowledge and clarity of preferences, why would there be such a plethora of books attempting to teach time-management, personal decision making, goal and priority setting, value clarification, etc. (a bibliography of time management books in Covey, Merrill et al. [1994] 2003 lists well over a hundred)?

While not supported, the unrealistic planning and perfect foresight of the original rational addiction model was difficult to overlook and hard to swallow. These aspects therefore became part of the motivation for Orphanides and Zervos (1995), whose theory of “Rational Addiction with Learning and Regret” stated that “the bulk of objections concerning earlier rational models can be attributed *not* to rational decision making, but rather to the common implicit assumption of perfect foresight” (740). The authors therefore introduce uncertainty through three “fundamental postulates”:

[T]hat consumption of the addictive good is not equally harmful to all individuals, that each individual possesses a subjective belief structure concerning his potential to become addicted, and that this belief structure is optimally updated with information gained through consumption, via a Bayesian learning process. (740)

Their model assumes that people are either vulnerable or invulnerable to harms from addictive goods. They would want to use drugs at low levels if they *knew* they were invulnerable, not at all if they *knew* they were vulnerable. To discover your type you have to consume drugs and note whether harms follow. Your prior beliefs about your type determine whether it is rational to experiment with low level consumption of drugs. If harm follows, the uncertainty is resolved. Since harms occur stochastically even for vulnerables, a vulnerable might have been consuming drugs for some time before he learns his true type. If he has consumed at a low level for a long time, the consumption stock has increased to a level making it optimal to implement an increase towards a high level of consumption and thus be an “addict.” If the harm happens after a short period of low level consumption it is rational to quit.

As an attempt to avoid the unrealistic foresight assumed in Becker and Murphy’s model this is a failure: The reason people do not wish to be addicted is now that they have calculated the present value of implementing the consumption paths we call addictive, and found these to have low present values compared to alternatives. In other words, merely to dislike addictions in this model requires the foresight and planning ability the model seemingly does away with. Also: Unstable consumption is still the implementation of a rational and detailed, forward-looking plan. Such plans are implemented when the uncertainty is resolved and the person discovers that he is a vulnerable. “I haven’t smoked much in the past but I had a smokers’ cough so bad my vision blurred this morning. Consequently, I plan to reach forty a day by the end of next month.”

While not denied, these forward-looking planning aspects of the model are not emphasized or supported by stories in the non-technical comments. The verbal discussion instead suggests the notion that people are afraid of addiction, the phenomenon left uninterpreted and perhaps seen as an undesirable illness. This, and the notion that people have a subjective estimate of their risk of becoming addicted, seems more plausible and intuitive to many. Philosopher Jon Elster, for instance, states that the model “captures the involuntary aspect of addiction that is central in real-life cases” (Elster 1999b, 185). The authors themselves seem to imply this interpretation when they write of people being “drawn into addiction” (Orphanides and Zervos 1995, 741), of addiction being “the unintended occasional outcome of experimenting with an addictive good” (741) and of addiction never being “the result of a deliberate plan but rather of what proves to have been an incorrect assessment of the possibility of becoming an addict” (751). This last one is interesting in that addictions *are* a deliberate plan in their model, addicts just rated the probability of this plan becoming optimal as low. Note also that the second “fundamental postulate” quoted states that people estimate their

“potential to become addicted,” whereas in the model they estimate their potential for experiencing harms. The difference is important, as we see when Orphanides and Zervos

[c]onsider an addiction to alcohol. To most people, current consumption of alcohol in moderation provides enjoyment without any serious side effects. To others, the same stream of consumption may lead to a dependence on alcohol, a harmful addiction. A predisposition to addiction, however, cannot be detected without the experience gained from repeated consumption of alcohol. (741)

Rewriting this in line with the model makes it not only less persuasive but counterintuitive: *The rational response to experiencing serious harms from moderate consumption of alcohol is often to plan and implement an increase in drinking. This detailed plan for future drinking can be called an “addiction,” and the longer you have been a normal, moderate drinker before you experience harm, the more likely it is that you will want to increase your consumption gradually towards addictive levels when, for instance, your liver gives out.*

Again, had the assumptions been correct, better stories would have been available. The theory claims that a) people consume large amounts *because* they experienced harms *from low amounts*, b) they increase their consumption only if the harms first appear after consuming at normal levels for a relatively long time, and c) those experiencing harms at earlier stages quit consuming. These implications should be easy to examine empirically. Especially claim a) seems to fly in the face of other views on addiction, where people experience harms *because* they consume at high levels, and are called addicted because they consume at these high levels *in spite of* the harm. Secondly, we would not have the phenomenon of “denial” where people deny that their high-level consumption causes them harm. Instead, people should say “of course I am experiencing problems from my consumption. Why else would I be consuming more and more?” Finally, as Elster (1999b, 183–185) notes, people should want to get and use information on risk factors for addictions (such as social background, personal characteristics, etc.) to improve the accuracy of their beliefs. This means that those with few risk factors should experiment more than those with many, other things equal. In other words, the factors predicting *use* should be the “opposite” of those predicting *addiction*.⁸

8. Since the number of “addicts” is the number of experimenters times risk of becoming an addict, the number of rational low risk users could become so much larger than the number

4.3. *A Simple, Stable, and Transparent World.* The planning involved in rational addiction theories requires beliefs of a kind we are not aware that we possess, and that we would not have evidence for if we did possess them. The beliefs involved are the kind required to pose the forward-looking decision problem on the mathematical form chosen by economists, and might therefore be viewed as a mathematical metaphor taken too literally. As Keynes put it,

Much economic theorizing to-day suffers, I think, because it attempts to apply highly precise and mathematical methods to material which is itself much too vague to support such treatment. (quoted in Chick 1998, 1864)

The danger is that the metaphor may be misleading—in our case by assuming precise, quantitatively specified relationships between variables—in ways that play too large a part in the mathematical argument—in our case enabling the identification of a precise, quantitatively specified consumption path that *makes no sense outside of the mathematical metaphor we have created.*

To see this, note that the mathematical solutions used by the theories require that the stock levels and investment equations are known: You can't design an optimal investment plan if your actions cause you to invest unknown amounts into an unknown asset with unknown rate of return that starts from an unknown level. This creates a number of problems.

First, since addictions may involve various drugs as well as “work, eating, music, television, their standard of living, other people, religion, and many other activities” (Becker and Murphy 1988, 676), constructing the stocks is hard: Anything with similar influences on welfare and incentives should be summarized by a single stock; e.g., if the death of a wife fills a husband with grief and increases the value of escape through drugs, this is a shock to the ‘addiction stock’ the husband could see as equivalent to a history of heroin use (Becker and Murphy 1988, 690–691).

Having identified the stocks, they should be operationalized, so that initial levels can be identified and changes monitored. E.g. what history of heroin use would equal the death of a spouse would equal what genes would equal what attitudes and expectations expressed in specified ways

of rational high risk users that this outweighed the difference in risk. Most addicts would then be from the low risk group, and if risk estimates were not conditional on use, this might be misinterpreted to suggest that high risk users were low risk users and vice versa. This could possibly generate “interesting dynamics.”

by various peers etc.? As managers say: If you can't measure it, you can't manage it.

Thirdly, beliefs in quantitative terms about the dynamics are required. E.g. a smoking teenager needs a precise estimate of how much and in what way both his next cigarette and the passing of time will influence his bodily organs, personal habits, future experience of smoking and withdrawal, relationships with various peers, family members, teachers, etc. He also needs such beliefs about future prices, tobacco policies, norms, etc.

Note especially the beliefs that should be formed concerning *how your tastes will change* in the future as a result of present experiences and choices. Conscious taste-planning is an important part of the motivation for rational addicts. Unfortunately, as shown in Loewenstein and Adler 1995, people fail to display awareness of even the "simple case" involving the sudden, robust, and highly systematic taste changes caused by acquiring an object ("the endowment effect").

Concerning the precision required of the beliefs: Plans of the kind discussed would be difficult with "qualitative beliefs" ("smoking damages future health") in place of numerically specified effects, since the trade-offs become unclear. Attempting to see qualitative beliefs as a more complicated form of quantitative beliefs (with a subjective probability distribution defined over different quantitative beliefs) will worsen the problem rather than avoid it (see Simon 1979, 505).

Even *assuming* quantitative beliefs, these will not be well-informed at the level of precision and detail required despite the fact that "some 70,000 scientific articles have implicated smoking in a wide variety of ailments, constituting the largest and best documented literature linking any behavior to disease in humans" (Chaloupka and Warner 2000, 1545). The reason is that it is still unclear how smoking patterns translate into health risks at the individual level. Smoking little is better than smoking lots (Leffondré, Abrahamowicz et al. 2002), but *reducing* your smoking may not reduce your risk (Godtfredsen, Holst et al. 2002). "[S]moking status, intensity, duration, cigarette-years, age at initiation, and time since cessation" are all possibly important variables in modeling the impact of smoking (Leffondré, Abrahamowicz et al. 2002, 813), but no consensus about their effects exist. Even the size and speed of health improvements for quitters is of uncertain magnitude, with persistence of elevated risk of coronary heart disease for ex-smokers being estimated at both 5 years and 30 years (Godtfredsen, Holst et al. 2002, 999).

Three further problems follow from this: The actual causal mechanisms will be far more complex than what rational addiction theories assume, and not just for tobacco. Skog (1999, 188–189) argues that this would also be the case for alcohol, and there is no reason to believe other

sets of causal mechanisms will be more accurately summarized by the simple and stable differential stock equations commonly assumed. Since forward looking planning is sensitive in that it generates different consumption patterns for different beliefs about causal mechanisms, this means that “predictive success” of rational addiction theory would only establish that real people act in a way that *would be* optimal if they were facing a problem we *know* they are not.

Second, the lack of well-known, detailed knowledge about effects makes it probable that different people will have different beliefs concerning the number of effects, their strength, lags, etc. This means that we might expect large behavioral differences due to unobservable differences in beliefs, which is problematic in that it becomes difficult or even impossible to generate potentially falsifiable predictions. Someone who expects such beliefs to vary a lot might even see similar behavior patterns across people and time as an anomaly for the theory.

Finally, the lack of precise knowledge means that the model is misspecified: Well-informed, rational people would *realize* their ignorance of the true structure of their decision problem. Their response should be to define a probability distribution over all possible specifications of the decision problem. The result might be a robust solution that does tolerably well across many specifications, taking into account that different plans lead to differing amounts of corrective feedback about the world. Orphanides and Zervos (1995) are on the right track in this regard, but their individuals know that their decision problem is one of only two possible. In reality, constructing the set of all possible specifications of the decision problem seems daunting, much less solving them.

4.4. Assume Away, Mr. Storyteller. The previous two sections concentrated on bringing out the mostly undiscussed and hidden assumptions made in the mathematics of rational addiction theories. The remaining question is why economists take these theories seriously as possibly valid explanations with a possible validity for welfare analysis. The answer seems to be that the theories seem consistent with the rational choice framework that economists are used to seeing, while presenting stories that give feelings of insight and that trigger feelings of understanding.

A good example to illustrate how these assumptions are supported concerns the necessary conditions for a harmful, rational addiction in Becker and Murphy 1988: The more you smoke, snort, inject, or otherwise consume of your addictive good today, the worse off you will be tomorrow (the level effect of the stock) but the more rewarding the addictive consumption will become (the higher the stock, the higher the marginal utility of the addictive commodity). Many stories can be used to illustrate these effects: They can be interpreted as withdrawal, lagged

health effects, etc. (Skog 1999, 175), they seem plausible, relevant, and important. The problem is that these stories are selected ad-hoc precisely for this reason, and that contradictory assumptions could be equally well supported by a different selection of stories.

For instance, we could tell stories that go against the necessary assumptions: A young boy can become so much better off through raised self-esteem, status, and a richer social network when smoking allows him to join the in-crowd at his school, that the level effect of smoking becomes positive and smoking becomes a beneficial addiction. A smoking adult can receive such pangs of remorse from each cigarette if he knows that smoking both disappoints his non-smoking friends, colleagues, and family members and makes them see him as a pathetic addict, that the effect of the stock on the marginal utility of cigarettes is reversed. According to the theory, an addiction would be impossible.

To disregard such stories simply because no *rational* individual would consume addictively unless the necessary conditions were fulfilled would be to beg the question by using the theory to constraint the stories about reality we accept.

The ease of telling stories suggests that almost any assumptions of qualitative effects can be supported by creative thought. For instance: Just as stories can be found implying that real addicts fail the necessary conditions for rational addictions, other stories imply that activities not usually considered addictive *may* satisfy these: Scratching an itch increases future itching (level effect), but also relief provided by each future scratching (marginal effect). The rational scratching addict maps out his lifetime scratching plan and scratches with increasing intensity until he reaches a high, stable level. If we add in a “skin soreness stock” we might support a theory of rational cyclical scratching. With stochastic effects and people with thick and thin skin we could get rational scratching addictions with learning and regret.

Of course, sometimes such stories are more easily found than at other times. Orphanides and Zervos (1995), for instance, lack good stories for the assumptions that there are invulnerables (or, more generally, initially invisible differences in vulnerability). Can some people smoke without risking lung cancer or bothering non-smokers? Get drunk without having their speech slurred, their thought confused and, with long term consumption, liver, and brain damaged? As we have seen, their verbal comments provide stories concerning differences in some unspecified addictive tendency rather than differences in harm vulnerability.

By telling stories we can also rid our theory of anomalies; e.g., old people do not need to worry about long term consequences. Since Becker and Murphy doubt the implication that the elderly will be addicts more often than others, further stories are introduced:

Of course, other things are not usually the same: older people are less healthy and subject to different life cycle events than younger people. Moreover, people who manage to become old are less likely to be strongly addicted to harmful goods. (Becker and Murphy 1988, 684)

Now, assume that the prediction had been wrong in the opposite direction: *Of course, old people have frail bodies that are poor at 'producing utility' from physical activities (sports, dancing, sex, etc.). They compensate by shifting towards mental experiences provided by drugs. Rather than lying alone waiting in vain for family visits, they relax in the soft, sweet daze of heroin or take dazzling trips on LSD.*

Finally, we may note that stories only provide qualitative effects and that the different stories we collect may apply to different individuals. Our end result may thus be a representative agent model that does not represent anyone, with whatever qualitative effects we desire, and with quantitative magnitudes of effects open for manipulation.

4.5. Who or What Are We Talking About? Many of the stories used to support the specification of rational addiction theories refer to intuition and common sense and presuppose that introspection can provide valid insights into how people make decisions, i.e. that we are aware of or can uncover the reasoning behind our choices. Since the assumptions concern *believed* consequences and *perceived* incentives, the use of introspection and common sense seems hard to avoid. Introspection, however, provides even stronger arguments *against* these theories: "TV-addicts," binge eaters, or chain smoking tobacco users hardly fit our idea of perfected, deliberate, conscious decision makers. We find it absurd to think they have mapped their future lives in all details. We have no awareness of precise, quantitative, and well-informed beliefs concerning the complex laws of nature and society required for stating the optimal control problem rational addiction theories claim that we solve. Add in that we should take these beliefs into account in a complicated way that would baffle anyone without extensive mathematical training, that we should do this even when the problem uniquely facing us is more complex than our formal mathematics can handle in practice, and all to yield plans that few if any of us are aware that we possess, and the theory becomes hard to take seriously. Especially when we recall that the whole, theoretical edifice has been raised to explain why people are lazy, fail to take care of their health, consume harmful substances that are pleasurable here and now, etc.

It is a paradox that the very theories claiming that addictions are well informed, rational, forward looking decisions clash strongly with how we perceive our own deliberate and conscious decision processes. This can

be avoided by denying that choices are made consciously. In an earlier approach Becker did this by stating explicitly that the decision problem could be formulated and solved at some not-clearly-identified and not-necessarily-conscious level, claiming support from psychology's belief in the subconscious (references in Rosenberg 1979, 513). As noted by Rosenberg such a defense of the theory means that "the ordinary interpretation of the theory which common-sense specification seems to suggest is ruled out" (Rosenberg 1979, 525).

In other words: If enjoyment^{Becker} has a different meaning from enjoyment^{common usage}, then the fact that someone *consciously* enjoys^{common usage} the effects of cannabis because it enhances their pleasure^{common usage} from food, sex or music has *no implication* for how cannabis affects their pleasure^{Becker} from food, sex or music and whether these effects are enjoyed^{Becker} or not. Why should even goods^{Becker} be assumed to involve the same categories as goods^{common usage}? The point is simple: When we begin referring to X by a name we use for something else, this does not teach us anything about X.

Denying that choices are conscious means that intuition, plausibility, and common sense now has *no value* for specifying the theory or justifying its assumptions. We need to specify what the decision making unit actually is, and what it maximizes: Do we want an extreme selfish gene theory that explains all behavior (including voluntary chastity, suicide, homosexuality, art, science, drug taking etc.) as purely attempts to maximize the spread of the organism's own genes?⁹ Do we want a Freudian theory where unobservable, postulated forces of Id, Ego, and Super-ego clash and conflict causing behavior richly symbolic of repressed childhood memories and sexual conflicts? Or do we postulate the unconscious as some guardian-angel/supercomputer that solves problems the consciousness faces but is unable to formulate and understand itself? How does it decide, plan, get information, update beliefs, etc.? And on *what possible evidence* could we base our assumptions about these matters?

Finally, if the decision maker is not the conscious person, how do we weigh the "welfare" (or "overall evaluation" if you will) of the deciding subconscious against the welfare of the deluded and passively experiencing along-for-the-ride consciousness? Should we care, in a policy context, about how "satisfied" the subconscious was with its attempt to maximize the spread of its genes? If we could thwart the subconscious in a way that made the conscious person feel more happy, satisfied and

9. This *extreme* selfish gene view is not intended as a representation of the views of Richard Dawkins, coiner of the phrase "selfish gene," who accepts that we may consciously choose actions that go against the "wishes" of our genes.

valuable, should we do it? This relates to the earlier discussion of whether the welfare claims can be valid when the assumptions are not: The theories of rational addiction are clearly not true at the conscious level. If they are true at some subconscious level, should we still care about the welfare analysis that flows from it? If the theory was correct whatever answer we give to these questions would actually serve our subconscious, so I'll leave the matter here.

5. Conclusion. I have here used rational addiction theory as an example of choice theories that are absurd when understood in the realist sense required by the claims made for them. These theories show how a loose, unstructured approach to explaining and justifying a mathematical model allows one to hide problematic assumptions even when these are central to the argument made, while providing ad-hoc illustrations that trigger feelings of understanding and insight though neither justifying the assumptions nor providing an adequate explanation in objectivistic terms.

Readers agreeing at this point that the theories are absurd might still wonder whether economists *really* take these theories seriously. My answer is split.

On the one hand: Yes. They take them seriously in their writings. They sell their articles by claiming real world relevance for their explanations and the derived welfare analysis. For instance, in their second addiction article Orphanides and Zervos conclude that their

analysis suggests that harmful addictive behavior is largely consistent with the standard axioms of rational, forward looking utility maximization and should no longer be considered a challenge to standard economic analysis. Rather, a remaining challenge is to employ standard economic analysis for further examination of the welfare implications of addictive consumption and provide better guidance for public policy design. (1998, 89)

On the other hand: No. I am personally unable to believe these economists *really* take such theories seriously. If they truly believed these theories held value for explanation and welfare analysis, then I believe they would have explicitly and in detail justified this surprising belief to convince skeptics. Until a good case for taking these theories seriously has been made, I believe it might make sense to stop seeing them as failed theories about the world. Instead, they could be seen as successful moves in a game where the goal is peacock displays of technical ingenuity and skill while explaining all behavior as the choices of well-informed, utility maximizing individuals, with higher marks the less plausible these assumptions would seem initially. One of the referees to this paper shared

a story that would fit such an interpretation, where the authors in their work only pretended to believe in rational choice theories because the rules of the game demanded it. Their work “solved” a puzzle the referee pointed out would puzzle only someone ignorant of well-established findings from psychology. “The authors of the paper—both smart economists doing very interesting work that spans disciplines—replied that they recognize their position is psychologically inaccurate and deliberately naïve, but they are trying to publish the paper in an economics journal, and the dominant assumptions in the field require it.”

One might of course feel that adults should be allowed to play whatever games they like as long as they are consensual. The problem with the choice theory game—if it is indeed only a game—is that it is played as though it was a scientific discipline doing research relevant to the real world. Apparently, the rules prohibit stating when you or others are just playing the game, thus blurring the boundaries between the game and the science of economics. The authors might know what they’re doing; the readers might not be in on the joke. Some of them get fooled by stories in theories loosely read or briefly referred to into thinking that economists have actually done valid research revealing, for instance, drug addicts to be rational and addictions unproblematic from a welfare perspective. Others write off large portions of the economics discipline—including its sensible parts—as nonsense. Both consequences are undesirable, and both could be avoided if economists would admit to playing theory-games if and when that is indeed what they are doing. Or, alternatively, if they could tell us why we should believe the surprising claims they keep making.

REFERENCES

- Adda, Jérôme, and Valérie Lechene (2001), *Smoking and Endogenous Mortality: Does Heterogeneity in Life Expectancy Explain Differences in Smoking Behaviour?* Discussion Paper 77, Department of Economics Oxford: University of Oxford.
- Becker, Gary S. (1996), “Preferences and Values”, in *Accounting for Taste*. Cambridge, MA: Harvard University Press, 3–23.
- Becker, Gary S., and Casey B. Mulligan (1997), “The Endogenous Determination of Time Preference”, *Quarterly Journal of Economics* 112(3): 729–758.
- Becker, Gary S., and Kevin M. Murphy (1988), “A Theory of Rational Addiction”, *Journal of Political Economy* 96(4): 675–700.
- Chaloupka, Frank J., and Kenneth E. Warner (2000), “The Economics of Smoking”, in *Handbook of Health Economics*, Amsterdam: Elsevier, 1539–1627.
- Chick, Victoria (1998), “On Knowing One’s Place: The Role of Formalism in Economics”, *Economic Journal* 108(November): 1859–1869.
- Covey, Stephen R., A. Roger Merrill, and Rebecca R. Merrill ([1994] 2003), *First Things First*, New York: Free Press.
- Cowen, Tyler (1989), “Are All Tastes Constant and Identical? A Critique of Stigler and Becker”, *Journal of Economic Behavior and Organization* 11(1): 127–135.

- Dawes, Robyn M. (1999), "A Message from Psychologists to Economists: Mere Predictability Doesn't Matter Like It Should (without a Good Story Appended to It)", *Journal of Economic Behavior and Organization* 39: 29–40.
- Dockner, Engelbert J., and Gustav Feichtinger (1993), "Cyclical Consumption Patterns and Rational Addiction", *American Economic Review* 83(1): 256–263.
- Elster, Jon (1997), "More Than Enough", review of *Accounting for Tastes*, by Gary S. Becker. *University of Chicago Law Review* 64(2): 749–764.
- (ed.) (1999a), *Addiction: Entries and Exits*. New York: Russell Sage Foundation.
- (1999b), *Strong Feelings: Emotion, Addiction, and Human Behavior*. Cambridge, MA: MIT Press.
- (2000a), *Alchemies of the Mind: Rationality and the Emotions*. Cambridge: Cambridge University Press.
- (2000b), *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints*. Cambridge: Cambridge University Press.
- Elster, Jon, and Ole-Jørgen Skog (eds.) (1999), *Getting Hooked: Rationality and Addiction*. Cambridge: Cambridge University Press.
- Fehr, Ernst, and Peter K. Zych (1998), "Do Addicts Behave Rationally?" *Scandinavian Journal of Economics* 100(3): 643–662.
- Feichtinger, Gustav, Cars H. Hommes, and Alexandra Milik (1997), "Chaotic Consumption Patterns in a Simple 2-D Addiction Model", *Economic Theory* 10(1): 147–173.
- Ferguson, Brian S. (2000), "Interpreting the Rational Addiction Model", *Health Economics* 9: 587–598.
- Friedman, Milton (1953), "The Methodology of Positive Economics", in *Essays in Positive Economics*. Chicago: University of Chicago Press, 3–43.
- Friedman, Milton, and Leonard J. Savage (1952), "The Expected-Utility Hypothesis and the Measurability of Utility", *The Journal of Political Economy* 60(6): 463–474.
- Gibbard, Allan, and Hal R. Varian (1978), "Economic Models", *The Journal of Philosophy* 75(11): 664–677.
- Godtfredsen, Nina S., Claus Holst, Eva Prescott, Jørgen Vestbo, and Merete Osler (2002), "Smoking Reduction, Smoking Cessation, and Mortality: A 16-Year Follow-up of 19,732 Men and Women from the Copenhagen Centre for Prospective Population Studies", *American Journal of Epidemiology* 156(11): 994–1001.
- Hausman, Daniel M. (1992), *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.
- Krugman, Paul (1998), "Two Cheers for Formalism", *Economic Journal* 108(November): 1829–1836.
- Kukla, Andre (1991), "Criteria of Rationality and the Problem of Logical Sloth", *Philosophy of Science* 58(3): 486–490.
- Laibson, David (2001), "A Cue-Theory of Consumption", *Quarterly Journal of Economics* 116(1): 81–119.
- Leffondré, Karen, Michal Abrahamowicz, Jack Siemiatycki, and Bernard Rachet (2002), "Modeling Smoking History: A Comparison of Different Approaches", *American Journal of Epidemiology* 156(9): 813–823.
- Levine, Harry G. (2000), review of "Getting Hooked: Rationality and Addiction", *Addiction* 95(3): 449–452.
- Loewenstein, George, and Daniel Adler (1995), "A Bias in the Prediction of Tastes", *Economic Journal* 105(431): 929–937.
- Miron, Jeffrey A., and Jeffrey Zwiebel (1995), "The Economic Case against Drug Prohibition", *The Journal of Economic Perspectives* 9(4): 175–192.
- Orphanides, Athanasios, and David Zervos (1995), "Rational Addiction with Learning and Regret", *Journal of Political Economy* 103(4): 739–758.
- (1998), "Myopia and Addictive Behaviour", *Economic Journal* 108(446): 75–91.
- Rogeberg, Ole (2003), "Preferences, Rationality and Welfare in Becker's Extended Utility Approach", *Rationality and Society* 15(3): 283–323.
- Rosenberg, Alexander (1979), "Can Economic Theory Explain Everything?" *Philosophy of Social Science* 9: 509–529.
- (1981), *Sociobiology and the Preemption of Social Science*, Oxford: Basil Blackwell.

- (1992), *Economics—Mathematical Politics or Science of Diminishing Returns*, Chicago: University of Chicago Press.
- Simon, Herbert A. (1979), “Rational Decision Making in Business Organizations”, *The American Economic Review* 69(4): 493–513.
- Skog, Ole-Jørgen (1999), “Rationality, Irrationality and Addiction- Notes on Becker and Murphy’s Theory of Addiction”, in J. Elster and O. J. Skog (eds.), *Getting Hooked*. Cambridge: Cambridge University Press, 173–207.
- Trout J.D. (2002), “Scientific Explanation and the Sense of Understanding”, *Philosophy of Science* 69: 212–233.
- Winston, Gordon C. (1980), “Addiction and Backsliding - a Theory of Compulsive Consumption”, *Journal of Economic Behavior and Organization* 1: 295–324.