

**Midbrain Mutiny: The
Picoeconomics and
Neuroeconomics of
Disordered Gambling**

Economic Theory and
Cognitive Science

Don Ross, Carla Sharp,
Rudy Vuchinich, and
David Spurrett

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

© 2008 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please e-mail special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Palatino by SNP Best-set Typesetter Ltd., Hong Kong, and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Midbrain mutiny : the picoeconomics and neuroeconomics of disordered gambling : economic theory and cognitive science / Don Ross . . . [et al].
p. ; cm.

Includes bibliographical references and index.

ISBN 978-0-262-18265-2 (hardcover : alk. paper)

1. Compulsive gambling. 2. Motivation (Psychology). 3. Reward (Psychology).
4. Choice (Psychology). 5. Neuroeconomics. 6. Economics. 7. Cognitive science.

I. Ross, Don, 1962-.

[DNLM: 1. Gambling. 2. Behavior, Addictive. 3. Cognitive Science.

4. Economics. WM 190 M627 2008]

RC569.5.G35M53 2008

616.85'841—dc22

2007034671

10 9 8 7 6 5 4 3 2 1

1

Is There Such a Thing as Addiction?

To some extent, everyone engages in the same activities repeatedly over time. Chaos would reign if we did not show such behavioral regularities. Much of this behavior is constructive and healthy, such as work, physical exercise, and social interactions, while some of it is destructive and unhealthy, such as excessive drinking, drug use, or gambling. Given current understanding, however, we cannot distinguish good from bad repetitious behavior simply by reference to type, since some people drink, use drugs, or gamble in moderation and therefore avoid the destructiveness of excess, while others work or exercise too much and turn an otherwise healthy repetitious pattern into an unhealthy one. A scientific issue of immediate relevance to the clinical professions is to develop a theoretical framework that allows a clear empirical demarcation of the boundary between normal/healthy and excessive/unhealthy repetitious behavior patterns. Although there is now no consensus on what form such a theory would take, most professionals would agree that excessive and unhealthy repetitious behavior patterns are characterized by (1) experiences of craving prior to engagement, (2) impaired control over engagement, and (3) repeated engagement despite negative consequences (Potenza 2006).

“Habit,” “addiction,” and “compulsion” are popular terms used by both lay persons and scientists to describe these phenomena. A crude gauge of the extent of their use is provided by the Internet, with Google searches for “habit,” “addiction,” and “compulsion” yielding 46, 48, and 6 million hits respectively. This indicates that repetitious behavior patterns are very interesting to a lot of people. This interest is certainly warranted, as the costs of unhealthy excessive behavior patterns are high by any reasonable standard. Alcohol abuse, drug abuse, and smoking annually cost the United States \$176 billion, \$114 billion, and \$137 billion respectively (Rice 1999), in terms of lost productivity, health

care costs, and damage from accidents. The indirect costs in terms of crime fueled by addiction—and in terms of emotional suffering, which is difficult to shadow price—must be many times larger.

Psychiatric syndrome definitions and classification systems provide a sort of official history of the development of terms used to describe repetitious behavior patterns. “Addiction” was professionally respectable in the mid-twentieth century, as it was included in the first edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*, or *DSM-I* (American Psychiatric Association 1952), as a diagnostic label under the more general category of “sociopathic personality disturbances.” The meaning of the term addiction was not defined, however, and the label applied only to “alcoholism” and “drug addiction” as subtypes. In 1957 the World Health Organization distinguished between drug “addiction” and “habituation” (World Health Organization 1957). An “addiction” was characterized by (1) an overpowering desire (compulsion) to continue taking the drug and to obtain it by any means, (2) a tendency to increase the dose, (3) a psychic and generally a physical dependence on the effects of the drug, and (4) a detrimental effect on the individual and on society. A “habituation,” on the other hand, was characterized by (1) a desire (but not compulsion) to continue taking the drug, (2) little or no tendency to increase the dose, (3) some degree of psychic dependence on the effect of the drug, and (4) detrimental effects primarily on the individual.

Subsequent classification systems moved away from use of the term “addiction.” A 1982 World Health Organization report recommended that the term “dependence” replace the term “addiction” (World Health Organization 1982). This recommendation was based on concern over the stigma arising from the negative social connotations of “addiction,” and on the hope that the term “dependence” would allow for more objectivity and precision in definition (Henningfield, Cohen, and Pickworth 1993). In keeping with this sentiment, *DSM-II* (American Psychiatric Association 1968) no longer used “addiction” as the overarching label for consumption disorders, but instead included “alcoholism” and “drug dependence” as separate diagnoses under the category of “personality disorders” and “certain other non-psychotic mental disorders.” “Alcohol addiction” was separate from “episodic excessive drinking” and “habitual excessive drinking,” with presumptive evidence (i.e., withdrawal symptoms) of physical dependence necessary for the “alcohol addiction” diagnosis. Interestingly, the description of the “drug dependence” diagnosis distinguished between addiction

and dependence, although the basis of the distinction was not explained. *DSM-II* contained no mention of gambling.

DSM-III (American Psychiatric Association 1980) included “substance use disorders” as its own overarching diagnostic category, and distinguished between “substance abuse” (pattern of pathological use, impairment in social or occupational functioning due to substance use, and minimal duration of one month) and “substance dependence” (presence of physiological dependence as manifested by tolerance and/or withdrawal), but did not use the term “addiction” to refer to either.

DSM-III featured “pathological gambling” (PG) for the first time, among “disorders of impulse control not elsewhere classified” along with “kleptomania,” “pyromania,” and “intermittent” and “isolated explosive disorders.” PG was defined as: “Gambling that compromises, disrupts, or damages family, personal, and vocational pursuits, as indicated by at least three of the following: (1) arrest for forgery, fraud, embezzlement, or income tax evasion due to attempts to obtain money for gambling, (2) default on debts or other financial responsibilities, (3) disrupted family or spouse relationship due to gambling, (4) borrowing of money from illegal sources (loan sharks), (5) inability to account for loss of money or to produce evidence of winning money, if this is claimed, (6) loss of work due to absenteeism in order to pursue gambling activity, and (7) necessity for another person to provide money to relieve a desperate financial situation.”

DSM-III-R (American Psychiatric Association 1987) provided a significant revision of the “substance use disorders” diagnostic criteria. Unlike *DSM-III*, evidence of tolerance or withdrawal was no longer required for a diagnosis of “psychoactive substance dependence.” Any three of the following nine symptoms was sufficient for a “dependence” diagnosis: (1) substance taken in larger amounts or for longer period than intended, (2) persistent desire or one or more unsuccessful efforts to cut down or control substance use, (3) a great deal of time spent obtaining or consuming the substance or recovering from its effects, (4) frequent intoxication or withdrawal when expected to fulfill major role obligations, (5) important social, occupational, or recreational activities given up because of use, (6) continued use despite recurrent problems related to use, (7) tolerance, (8) withdrawal, and (9) use in order to avoid withdrawal.

In *DSM-IV* (American Psychiatric Association 2000), the current diagnostic system, the criteria for “substance-related disorders”

remained essentially unchanged from *DSM-III-R*. Again, PG was included among “disorders of impulse control not elsewhere classified,” along with “kleptomania,” “pyromania,” and “intermittent and isolated explosive disorders.” The diagnostic criteria for PG (see chapter 2) were modified to focus less on financial/criminal aspects and more on psychological/behavioral aspects.

Two general points are relevant regarding these diagnostic systems. First, the evolution of *DSM* demonstrates how syndrome definitions, classification systems, and criteria are not pure scientific concepts, and do not develop on the basis of new scientific evidence alone. Rather, these systems and criteria are heavily influenced by wider cultural, political, and economic forces. Babor (1990, p. 33) notes that “The primary issue with regard to definitions of dependence [or addiction] thus becomes who or which group controls the defining process, and how they use definitions to promote their own ends, be they medical, legal, scientific, or moral.”

Second, there is the question of whether one can explain a person’s excessive engagement in an activity by appealing to their “addiction” or “dependence.” Historically, the term “addiction” has been used to attempt to explain the excessive consumption of alcohol for well over 300 years (e.g., Warner 1994), of opiates and other now-illicit drugs for about 100 years (e.g., Levine 1978), of tobacco for several decades (e.g., Henningfield, Cohen, and Pickworth 1993), and, more recently, of gambling (Marks 1990). British clergy in the early seventeenth century initiated the use of the term “addiction” to describe substance use in reference to excessive alcohol consumption (Warner 1994). At that time, “addiction” meant an obligation or devotion to something. Thus, the term “addiction” became a metaphor describing a person’s pathological preference for a given activity, which leads him or her to behave “as if addicted” (i.e., obligated) to the activity. There is a standing tendency for metaphors of this sort to be reified, internalized, and attributed causal significance (Sarbin 1968). This appears to have happened over the years with the term “addiction,” which has gone from being a metaphorical description of devotion to a particular activity, to being an hypothesized internal state, the existence of which is revealed by that devotion, and, finally, to being the cause of that devotion.

Definitions and diagnostic criteria for addiction or dependence have generally been descriptive and not explanatory (Akers 1991); that is, they have denoted behavioral, cognitive, and emotional patterns and

associated consequences that have proven useful in reliably placing individuals into categories based on similarities. The traditional diagnostic criteria and category labels do not explain why individuals engage in those patterns, however. It may be a useful description to say that an individual's excessive engagement meets the diagnostic criteria for addiction or dependence, or that some activity is addictive, but this does not mean that it is useful to say that an individual's excessive engagement occurs *because* of addiction or dependence. Akers (1991, p. 779) makes this point as follows:

The problem is that there is no independent way to confirm that the addict cannot help himself, and therefore the label is often used as a tautological explanation of the addiction. The habit is called an addiction because it is not under control, but there is no way to distinguish between a habit that is uncontrollable from one that is simply not controlled. However precisely it is defined, addiction is a label, a term applied to behavior. It cannot, itself, provide an explanation for that behavior. We label excessive involvement in drugs which a person cannot seem to give up as addictions, and we think we have explained the excessive, hard-to-stop behavior by saying that the person is suffering from an addiction. The label of addiction is attached to the behavior because the person is assumed to have lost control of the substance use; when asked why the person has lost control, the answer is that he is addicted. In other words, addiction causes addiction.

A key issue is whether activity engagement becomes controlled by something other than the individual's choice to continue that use. Advocates of explaining excessive engagement by appealing to the individual's "addiction" believe that in addicts use becomes something other than a choice. For example, the 1988 Surgeon General's report (US Department of Health and Human Services 1988, pp. 248–249) asserted that "the term 'drug dependence' or 'drug addiction' refers to self-administration of a psychoactive drug in a manner that demonstrates that the drug controls or strongly influences behavior. In other words, the individual is no longer entirely free to use or not use the substance." Alterations in some aspect of brain function as a result of substance use typically are thought to override the individual's choice. This argument is made forcefully by Henningfield, Schuh, and Jarvik (1995, p. 1715) in relation to tobacco use: "The pathophysiological consequences of tobacco smoke exposure include tissue destruction contributing to lung disease, cellular changes contributing to cancer, and the cellular and molecular reinforcing effects leading to dependence. Once the pathophysiological consequences of tobacco use have

occurred, it may be no more a matter of personal choice to abstain from tobacco than to reverse metastasizing lung cells."

An alternative view was expressed by Heyman (1998, p. 807), who argued that "neuroadaptation could just as likely influence preference as preclude it. The difference is important. An addict who takes drugs voluntarily can be persuaded by contingencies or new information to stop using them. An addict who takes drugs involuntarily cannot be persuaded by costs and incentives to stop using them." Studies have shown systematic inverse relations between price and consumption of a variety of "addictive" substances, and between the availability of alternative activities and substance consumption (see Vuchinich and Heather 2003 for an overview). Such work by behavioral economists builds a picture of excessive engagement as the outcome of a reinforcement or utility maximization process. In this view, the activity is chosen to the extent that the reinforcement or utility gained from it exceeds the reinforcement or utility gained from engagement in other activities. Importantly, models of excessive consumption (i.e., "addiction") in this literature (e.g., Becker and Murphy 1988; Heyman 1996; Rachlin 1997) describe how substance use over time could change the utility or reinforcement gained from that use relative to the utility or reinforcement gained from other sources, rendering the former relatively more attractive. Thus, according to many behavioral economists, an individual's current choice to use a substance is influenced by their prior consumption, but both now and in the past the individual is choosing to consume the substance. The implications of this literature are that "(i) drug use in addicts can be altered by the proper arrangement of costs and benefits, (ii) addictive drugs reduce options but do not eliminate choice, and (iii) the biology of addiction is the biology of voluntary behavior" (Heyman 1998, p. 808). It is important to point out that in the behavioral economics (BE) literature, concepts of "voluntariness" and "choice" are not used so as to take a side in philosophers' debates over the existence of free will, or to morally convict excessive consumers of bad character. Rather, the intended emphasis is on the way in which the classical "addictions" are sensitive to positive and negative incentives in the same sense as non-excessive consumption; thus, they are cast as fit explananda for economics.

In the chapters to come, we will argue that current science appears to vindicate the insights of *both* behavioral economists and proponents of internal state (or, less accurately, "disease") models of addiction, dissolving the traditional tension between these views. We regard this

as a surprising development. Unlike the dominant intellectual tradition in psychiatry, which prefers to trace the roots of all pathology to a mechanical or chemical condition in the brain, our thinking begins at the molar scale, in economics. From this perspective, prospects for reduction to strictly molecular scales look *prima facie* doubtful; however, the view to which we have been led by consideration of empirical evidence fits neither the neat model of addiction as an endogenous disease nor the clinically defeatist model of addiction as an irredeemably complex social syndrome enmeshed in so many layers of cultural and moral construction that it is better handed over from scientists to novelists and oral historians. We are convinced that addiction is complicated but comprehensible, subject to some robust generalizations, and potentially much more effectively treatable than is usually assumed in popular discussions.

Let us first say a bit more about the complications our perspective urges on psychiatry. Given the growing volume of basic and clinical research using neuroscientific tools to study psychiatric disorders, one might naively assume that psychiatry is in the process of transforming itself through full and fearless exposure to these new investigative technologies. In fact, strong methodological and conceptual conservatism largely continues to hold sway. For the most part, only conditions that psychiatrists have already modeled as “organic disorders” are investigated under the rubric of neuropsychiatry, so their defining properties are treated as fixed in advance of the new research. Organic disorders are individuated by their “psychiatric presentations” (Agrawal 2004)—that is, their manifestations in degenerative disorders like dementia, head injury, strokes, or epilepsy, in movement disorders like Parkinson’s disease, intracranial tumors or infections, and in nutritional, toxic or endocrine disorders. It is only more recently that psychiatric disorders such as schizophrenia and autism have tentatively been added to the investigative domain of neuropsychiatry. In the 1960s autism was thought to be caused by a “refrigerator mother.” Evidence in the last two decades now provides overwhelming support for the notion that autism is very much a brain disorder.

Historically, the division of psychiatric disorders into organic and non-organic types was based partly on the fear that reducing them to underlying brain pathologies would imply disregard for environmental influences. In most of the behavioral sciences, decades of heated “nature–nurture” controversy have given way to a general acknowledgment of bidirectional feedback relationships between biological and

environmental influences as the norm in behavioral ontogenesis and regulation. For instance, thanks to the successful integration of neuroscience into parenting research, we now understand how sensitive caregiving affects the developing brain of an infant (Sharp and Fonagy forthcoming; Fonagy et al. 2002). As an example of causal influence in the other direction, it has recently been appreciated that reduced amygdala functioning in psychopaths affects their ability to recognize fear in others, leading to behavior that is inadequately governed by societal norms (Blair 2001, 2003; Blair et al. 2006).

In this general context of opening new windows between psychiatry and neuroscience, our perspective from an economics standpoint leads us to ask how neuroeconomics (NE) can best contribute. In chapter 5 we will make extended use of the new neuroeconomic model of valuation in the brain's so-called reward circuit. Valuation is the process by which reinforcement signals provide feedback to an organism regarding the predicted value of a stimulus. To cite an evocative example, the mysterious pleasure one feels when one "clicks" with another person (a "meeting of minds"), a key contributor to the sense of connectedness humans share, can be captured by a positive valuation assigned to the experience by the brain. As the neuroeconomist Read Montague puts it in his recent popular book (2006, p. 20), "valuations put meaning back into computations."

It is also the case, however, that many humans do not receive the same positive reinforcement signals from stimuli that typical people do. For instance, some individuals, notably psychopaths, rarely experience a sense of connectedness or a meeting of minds. Psychopathy is characterized by a combination of antisocial behavior, proneness to boredom, impulsivity, a callous interpersonal style, superficial charm, and a diminished capacity for remorse (Cleckley 1941; Hare 1991). Within a reinforcement learning framework, experimental paradigms derived from NE, such as economic exchange games, can be usefully applied to understand psychiatric disorders such as psychopathy. For example, Rilling et al. (2007) used an iterated version of the prisoner's dilemma to probe the neural correlates of interpersonal deficits associated with psychopathy. Results showed that subjects scoring higher on psychopathy, particularly males, defected more often and were less likely to continue cooperating after establishing mutual cooperation with a partner. Further, they experienced more outcomes in which their cooperation was not reciprocated. After such outcomes, subjects scoring high in psychopathy showed less amygdala activation, suggesting

weaker aversive conditioning to those outcomes. Compared with low-psychopathy subjects, subjects higher in psychopathy also showed weaker activation within the orbitofrontal cortex when choosing to cooperate and showed weaker activation within the dorsolateral prefrontal and rostral anterior cingulate cortices when choosing to defect.

By applying NE and computational models of reinforcement learning to disorders such as psychopathy, in the same way that we will do in this book to PG, we may begin to describe how computational valuation may go awry in a variety of different psychiatric disorders. While a significant literature exists that applies computational neuroscience approaches to probe psychiatric disorders (see Williams and Dayan 2005; Williams and Taylor 2004; Dayan and Williams 2006), it is only recently that studies are conducted under the umbrella of an emerging new field called computational psychiatry (Dayan and Williams 2006; Montague 2006). This is founded on the fact that most psychiatric disorders involve neuromodulators and that an established literature exists on the normative function of these neuromodulators through work in computational neuroscience.

In retrospect, it is clear that one of the half-dozen most momentous scientific breakthroughs of the past century was McCulloch and Pitts's 1943 demonstration that neurons are computers, transformers of informational input into informational output, in a precise and generalizable mathematical sense. This insight was extensively developed by artificial intelligence researchers and philosophers of cognitive science during the succeeding decades, in consequence of which it gradually became clear that the mysterious relationship between mind and brain had at last been decisively illuminated by the concept of computation. That concept allows us to understand how the mind can be more than a metaphorical residue of nonphysical Cartesian spirit, while nevertheless not simply being equivalent to the anatomical matter of the brain. Like software that is "real enough" to make its designers wealthy, the mind exists "virtually," as patterns of information processing and computations that the brain specifically implements in the electrodynamics and molecular chemistry of its enormous neural network. Amit (1989) and Churchland and Sejnowski (1992) mark the earliest consolidation of the mature conceptual synthesis.

Cognitive information processing initially modeled in the abstract frameworks of artificial intelligence (e.g., universal Turing machines, von Neumann architectures, or generic parallel distributed processing

nets) is now increasingly being grounded in biological information processing. The detailed interanimation of molecular biology and cognitive science is putting testable, exact structure on the idea, articulated decades ago by a few philosophers such as Daniel Dennett (1969), that the algorithms of mental function (in philosophical jargon, the inferential patterns constituting “intentionality”) are based on brain structure, as interpreted in the context of social interaction. Montague (2006) now articulates the significance of these developments for neuropsychiatrists. “Structure and function,” he tells them, “are really information processing being implemented by the physical and chemical properties made available by biological molecules, cells, networks of cells, and so forth” (Montague 2006, p. 15).

Despite our appreciation of the importance of neurocomputational modeling, our perspective remains fundamentally behaviorist.¹ Contrary to stereotype, behaviorists need not and should not deny that brains perform computational processes or generate representational content for other brain processes to use as inputs. Nor need they or should they deny that (many) animals have subjective experiences: some of the representations that brains produce are accessible to second-order awareness. Sensible behaviorism as we understand and endorse it consists in the following philosophical convictions (explained and defended at length in Ross 2005, chapter 1):

(1) Computational models (including neurocomputational models) of the brain are functionally parsed by reference to the contributions of specific modules and processes to behavior. Of course, neuroanatomy puts strong constraints on such parsing; if a computational model says that functional module *a* is realized in neural area *x*, that functional module *b* is realized in neural area *y*, and that *a* sends output to *b*, then either *x* and *y* had better turn out to be synaptically connected in the right direction, or *x* had better produce changes in neurotransmitter levels that demonstrably modulate the activity of *y*. But to say that, for example, cells in the superior colliculus process retinal information is to say that those cells make a contribution of a certain sort to the behavior of extracting and using information about the distribution of light in the ambient environment—that is, seeing.

(2) Behavioral patterns are not individuated by reference to computational processes or representations in the brain. They are individuated by reference to the roles they play in organisms’ ecological interactions. In the case of the peculiar hyper-social mammal *H. sapiens*, the over-

whelmingly most salient such interactions are social, and are culturally mediated (see Sterelny 2003 for an account of cultural mediation); thus, what counts as behavioral pattern p in one ecological context might not count as another instance of p in a different ecological context. What we are affirming here is the view philosophers call “externalism about mental content,” which is the dominant perspective among philosophers of mind at present. Philosophers defending this view whose names are most likely to be familiar to behavioral scientists include Ludwig Wittgenstein, W. V. Quine, Wilfrid Sellars, Gilbert Ryle, Daniel Dennett, and Andy Clark.

(3) Because functional brain processes and types of representations are (ultimately) individuated by reference to their roles in behavior, and because types of behavior are individuated ecologically, it would be very surprising if brain processes and types of representations individuated only by reference to neurophysiological properties lined up neatly with functional brain processes and types of representation. That outcome would imply extraordinary success by past behavioral sciences in inferring invisible brain structures entirely on the basis of observing complex behavioral patterns.

(4) Therefore, in general, neuroscientific explanations of psychological processes will take the following logical form. A neuroanatomical–neurophysiological–neurochemical–neurodynamic account will explain a neurocomputational account. The neuroanatomical (etc.) facts that do the explaining will use different—noncognitive—concepts from the neurocomputational account. The neurocomputational account will in turn combine behavioral and nonbehavioral concepts to partly (sometimes mainly) explain psychological phenomena described in behavioral terms. Full explanations of behavioral regularities will more often combine neurocomputational, developmental, social, and cultural facts.

Among the behavioral sciences, economics begins at further logical distance from the brain than any other. This is because, its practitioners’ frequent commitment to *normative* individualism notwithstanding, economics since the mid-1930s has been mainly concerned with *aggregate* social phenomena—for example, constructing community indifference curves for representing aggregate demand. These curves cannot be decomposed into the behavior of individual members of the community in question, let alone into contributions of the brains of individuals. If the terms in which microeconomic phenomena were

individuated mapped neatly onto the terms in which neurochemical processes are described, this would be doubly miraculous, since it would rely on an improbable mapping from the microeconomic to the psychological, and then on another improbable mapping from the psychological to the neurochemical.²

Despite this, we are not skeptics about the thriving science of BE and its exciting new collaborator NE (Glimcher 2003; Montague and Berns 2002). From the fact that microeconomic processes do not *reduce* to psychological ones, it does not follow that psychology isn't relevant to explaining some microeconomic regularities. Similarly, from the fact that behavioral processes don't reduce to neurocomputational ones, it doesn't follow that neuroscience isn't relevant to explaining behavior. (In the second case, if this *did* follow, it would constitute a *reductio ad absurdum* of the antireductionist claim. We expect neuroscience to swamp all other sources of behavioral explanation in importance; we just don't expect it to completely *replace* them.)

NE finds useful purchase because the brain, like an economy, is a parallel processor of information in which units compete and cooperate for scarce resources. Viewed "from the bottom up," the scarce resource for which neurons compete is blood hemoglobin. Functional magnetic resonance imaging (fMRI) monitors the tussle over this asset. Viewed "from the top down," the scarce resource for which neurons compete is influence over behavior. (This is the "neural Darwinian" perspective of Edelman 1987.) We do not expect BE (or, to anticipate ideas we will introduce in chapter 3, "picoeconomics") to reduce to NE (Ross 2005, forthcoming), but we expect a steady parade of fruitful insights as the top-down and bottom-up perspectives are adjusted to accommodate and shed light on one another—that is, as NE partly explains picoeconomic phenomena in neurochemical and neurodynamic terms, and picoeconomics (PE) partly explains neuroeconomic phenomena in functional-behavioral terms.

Now let us come back to addiction. As we reviewed above, the history of this concept is rooted in behavioral observation, and scientifically shallow observation at that; thus, it might have been expected that, at best, addiction would not project usefully into the brain or, at worst, addiction would not have even stood the test of rigor as an organizing concept for behavioral science (including BE). No philosophical presupposition is ever a guarantor of scientific experience, however. In any case, philosophical externalism does not deny that

ecologically centered observation does not *sometimes* detect underlying internal mechanisms and then successfully track them across ecological contexts. (So, for example, we take it that popular observation picked up on strokes and successfully tracked them; on the other hand, it was not so successful at distinguishing the internal correlates of anxiety and depression.) We think that recent neuroscience strongly suggests that addiction is an internally governed kind of state that, in addicts, sufficiently dominates ecological variables as to be behaviorally stable and salient across ecological (including social) contexts. We argue that this is particularly evident when one understands addictive responses by reference to the neuroeconomic model of the dopaminergic reward system.

This general thesis emerges from our *assuming* a strong working hypothesis. This hypothesis is that if there is a core manifestation of addiction at the neural level of analysis, this should be gambling addiction. Behaviorally, as we will discuss in the next chapter, pathological gamblers resemble substance abusers very closely. We would expect (rightly, as it turns out) that *if* there is a common neural process underlying all behavioral addictions, this will be accompanied by different neurochemical aspects in the case of different drugs on which people get hooked. These differences are likely to constitute noise through which the hypothetical core neural process of addiction is obscured. By contrast, *if* there really is a distinctive neural phenomenon of addiction, and *if* some people who gamble more than they aim to are addicted to gambling, then gambling is an ideal candidate to serve as the basic model of addiction in general. If there are gambling addicts, and if it is a neurochemical process of a certain sort that makes these individuals so, then in gambling addiction the brain develops the pathology using only endogenous chemical resources. If one could identify these, the obvious next step would be to examine how exogenously introduced chemicals on which people become dependent bring about, by multiple pathways, the same endogenous disorder displayed in shining causal isolation in the case of gambling addicts.

We will argue in chapters 6 and 7 that our hypothesis above is well supported by the evidence; thus, where addictive gambling is concerned, our position will closely match the expectations of the reductive neuroeconomist. We doubt very much, however, that all people who have gambling problems according to *DSM* criteria are gambling addicts. The problem of the non-addicts, insofar as they *share* a common problem, is ecological rather than neurological. We do not thereby

mean to suggest that non-addicted problem gambling behavior is causally independent of all brain processes; no behavior is, of course. Our point is that non-addicted problem gamblers may not—probably do not—share any *common* neurochemical syndrome that explains what they have in common behaviorally. For understanding and treatment of the behavior of people falling into this category we will recommend appeal to a piceoeconomic model.

This book thus constitutes, as indicated in the “Note to Readers about *Economic Theory and Cognitive Science*,” an application to a particular phenomenon of Ross’s hybrid piceoeconomic–neuroeconomic framework for relating economic theory to the cognitive and behavioral sciences. At the same time we intend the work to stand on its own as an account of addiction and of the varieties of gambling disorders, demanding neither understanding of nor agreement with Ross’s general explanatory program. (Thus, for example, a reader might decide that the program works well enough in application to disordered gambling, but fails with respect to other phenomena.)

The book is organized as follows. Chapter 2 provides a general introduction to the scientific study of disordered gambling (DG). This study has naturally been motivated chiefly by clinical and policy concerns, which crucially interact with basic scientific issues throughout the book. In this respect, DG, and impulsive consumption in general, resemble all problems in applied economics. Like most scientific disciplines, economics is partly driven by philosophical curiosity and partly by our needs for successful engineering.

We think it useful to remind ourselves, and readers, of this practical side of the issue before chapter 3 wades into foundational problems around the *kind* of theory that best captures both scientific and clinical objectives. The chapter first presents our view of the relationship between standard (“neoclassical”) microeconomic theory and BE. In particular, it explains why we view the latter as complementary to the former rather than as a challenge to it (as much recently fashionable literature urges). We then explain the foundations of, and principal models and variations in, the subdivision of BE that we apply to DG. George Ainslie has dubbed this subfamily “piceoeconomics,” and this is usage we embrace in the subtitle of the book. The piceoeconomic perspective explains impulsive consumption in such a way as to shift the principal puzzle before us from the question “Why do some people suffer from chronic consumption impulsivity?” to the question “How do most people avoid such behav-

ior?" We explain the principal mechanism to which PE appeals in addressing this second question: establishment of "personal rules" as equilibria in negotiations among virtual subpersonal interests with divergent preferences over reward scheduling and reward types. Chapter 3 concludes with a review of leading criticisms of the picoeconomic model that have recently been articulated by other scientists. We provide partial responses to some of these, completion of which is deferred until the end of the book following mustering of relevant empirical evidence.

Chapter 4 begins that review. It surveys experimental investigations, first into consumption impulsivity in general and then into DG in particular, that test the predictions of the picoeconomic model. As we will see, the model has proven to be a fruitful organizing framework for designing empirical research and consolidating its results; this cannot be regarded as having proven conclusive, however, and we believe there is a principled reason for this. Most disordered gamblers do not *merely* have difficulty establishing and maintaining personal rules, though such difficulties are indeed a necessary condition for their problem. Subjects recruited for comparative studies of disordered gamblers and controls are selected using instruments, discussed in chapter 2, that attempt to diagnose people who are thought to suffer from a psychiatric pathology of the old-fashioned sort as characterized earlier in this chapter. Evidence strongly indicates that PE, as a molar-scale account of behavioral patterns, is a partial and essential, but incomplete, account of what goes wrong for severely disordered gamblers—so-called "pathological gamblers." To describe the situation in terms of the concepts discussed above, pathological gamblers indeed suffer from what traditionally minded psychiatrists might have called an "organic disorder." This neuropsychiatric condition, we maintain, has empirically turned out to meet the core criteria for true addiction, and conceptual confusion is best minimized if we regard addiction as reduced to it.

Though addiction may be "organic" in the sense of being identifiable by neuroscientific (in this case, neurochemical) properties, it is a neurocomputational model that allows us to identify it and systematically connect these properties to the behavioral patterns classically regarded as addictive. More specifically, the model is neuroeconomic. This newest subdiscipline of neuroscience, which appeals to economic theory and modeling techniques to understand the relationship between valuation and behavioral control at the scale of brain processes,

is given a general introduction in chapter 5. That chapter then presents the structure of and empirical evidence for the emerging neuroeconomic model of addiction as a distinctive pathology afflicting a functionally and chemically specifiable part of the brain known as the dopamine reward system. In consequence (probably) of interacting contributions from genetically inherited vulnerability and developmental contingencies, this system can usurp control of the motivational, attentional, and even cognitive and conscious aspects of the whole person. In effect, this midbrain circuit commits mutiny against the normal personal control apparatus. For reasons explained by the neuroeconomic model, the dopamine mutineer, considered as an economic agent, then maximizes its utility by relentlessly pursuing goods with certain properties that the targets of addiction all share.

It is often supposed that these properties must be, or must reside in, exogenous chemicals that addicts introduce into their bodies by ingesting substances. The best evidence that this is not the case comes from empirical application of the neuroeconomic model of addiction to PG. In chapter 6 we review this empirical work, pausing at various stages to consolidate its theoretical significance. PG, we conclude, is not only genuine addiction in the neuropsychiatric sense that previous parts of the book have identified; it is *the* variety of addiction which, by involving minimal incidental effects on other brain processes of the sort brought about by exogenous chemicals, provides the cleanest window on addiction as an endogenous neuroeconomic–neurochemical phenomenon. Simply put, pathological gambling is the basic form of addiction, the form on which drug addictions are then special complications for purposes of general understanding.

Only part of the evidence for this involves what philosophers of science think of as “theory (or model) testing.” We would not be nearly so confident in our conclusion were it not for the fact that, over the past three years, natural neuropharmacological experiments with psychiatric patients—that is, medical accidents—have exposed control levers on PG that can be manipulated by drugs. The drugs in question operate on exactly the part of the brain the neuroeconomic model would predict, and in the way it would predict. Science is ultimately about such deliberate manipulation of real phenomena, its concepts, arguments, deductions, generalizations, models, and theories all being devices to that end, and evaluated (by scientists, if often not by philosophers) in that light.

Because of the importance we attach to this kind of evidence, we devote chapter 7 to it. This is also where we again remind readers that the scientific problems involved cannot be detached from the underlying clinical issues. Chapter 7 therefore considers nonpharmacological treatments of DG in addition to drug therapies. In the course of this, we begin to build our case for merely *local* reduction of *pathological* gambling to addictive gambling, in a wider nonreductive framework that continues to model *disordered* gambling in general as a behavioral (molar-scale) syndrome. This outcome, which philosophers who like their reductionism all-or-nothing might find inelegant, is, we suggest, typical of scientific progress in the conceptually messy world that is governed by chance and contingency, not logical order.

The concluding chapter 8 is devoted to several jobs of consolidation. First it takes seriously the clinical imperative, asking whether our proposed local reduction of PG to addiction threatens to inflate the concept of addiction in such a way as to encourage confused medical policies. As we point out, this is a worry to which we must be particularly alert if, as we argue, addiction will soon be most appropriately treated with powerful, and hence potentially dangerous, neurochemical agents. The most serious ground for such concern, we show, is the recent appearance of an industry built around “diagnosing” and “treating” a non-scientific idea of addiction to sex. Since this industry is mainly a part of an ideological campaign rather than a genuine medical one, its expropriation of a concept that is, according to us, neuroscientific in its basis should be resisted on both scientific and welfare grounds. We then argue that local reduction of addiction by the neuro-economic model is likely to make such resistance more rather than less effective.

This returns us to more narrowly scientific challenges to our combination of local reductionism and global antireductionism regarding addiction. We conclude the chapter, and the book as a whole, by returning to these themes, which had been left partially argued at the end of chapter 3. Neuroeconomics and its investigative tools are so novel and liberating, the new paths to knowledge that they open so exciting, that it is only natural for scientists to want to push their application as far as it can possibly go. This should be encouraged, we believe; however, we reiterate and expand on our reasons for belief that PE, and therefore BE and standard microeconomics more generally, will continue to play essential roles in both generalizing our understanding and guiding clinical practice. Our overall model of DG, finally consolidated and

summarized in chapter 8, is a hybrid picoeconomic–neuroeconomic model of the kind characterized by Ross (2005, 2007, forthcoming). Of course, we are in a position here to promote this hybrid only with respect to our particular subject of study—DG—but since we argue that DG is the basic form of addiction, and that addiction in turn is the form of reward system breakdown that best reveals the nature of the neuroeconomic mechanism, we take our conclusion to have quite general significance.