Syndrome stabilization in psychiatry: pathological gambling as a case study

Don Ross†

†4th Floor, Humanities Building
University of Alabama at Birmingham
900 13th Street South
Birmingham, AL
USA 35294-1260

†School of Economics
Leslie Social Science Building
University of Cape Town
Private bag
RONDEBOSCH 7701
Cape Town
SOUTH AFRICA

don.ross@uct.ac.za
ABSTRACT

Murphy (2006) criticizes psychiatric nosology from the perspective of the philosophy of science, arguing that the model of pathology as encapsulated in the *Diagnostic and Statistical Manual of Mental Disorders* reflects a folk conception of the mental, and of malfunctioning, that is inadequately integrated with cognitive and behavioral neuroscience. The present paper supports this view through a case study of research on pathological gambling. It argues that recent modeling based on fMRI studies and behavioral genetics suggests a stipulative, non-seamless reduction of pathological gambling to a specific disorder of the mesolimbic dopamine system. This argument is agnostic as between prior philosophical commitments to realism or empiricism.

Word count, abstract: 106

Word count, text (including references and notes): 4,875
1. Introduction

Systematic research into gambling behavior, especially problem and pathological gambling, has become a substantial academic industry. Several research centers and an annual conference draw investigators from psychiatry, psychology, neuroscience, and social sciences (particularly economics). Historical moments when interdisciplinary research matrices converge on norms that guide grant and journal reviewers are good opportunities for study by philosophers of science. They are also the occasions when philosophers of science can most usefully make contributions to science, since they are the junctures at which the relevant scientists will often agree that conceptual uncertainties merit self-conscious attention. The current state of gambling behavior research exemplifies this situation, while also revealing features specific to conceptual stabilization at the intersection among psychiatry, neuroscience and social science.

Murphy (2006) criticizes standard practice in psychiatric nosology as a philosopher of science. He argues that the model of pathology encapsulated in the *Diagnostic and Statistical Manual of Mental Disorders* IV (henceforth *DSM-IV*) (American Psychiatric Association 2000) reflects a folk conception of the mental, and of malfunctioning, that is inadequately integrated with cognitive and behavioral neuroscience. The present paper
supports this view in the specific instance of pathological gambling.\textsuperscript{1} The case does not rely on philosophical presuppositions of realism or empiricism, either of which, if dogmatically conventionalized in a science, would unduly constrain opportunism. However, realism and empiricism are useful constructs for describing particular trade-offs that scientists make between the motive to find unifying explanatory (causal) mechanisms (realism) and exploitation of statistical testing power provided by reduced-form models that are agnostic about constituents of model-independent reality (empiricism).

2. The \textit{DSM} operationalization of pathological gambling

From clinical lore gambling research inherited distinct constructs of ‘problem’ and ‘pathological’ gambling. This classification was motivated by reference to social criteria: some people’s gambling is widely deemed to be \textit{generally} socially catastrophic (to gamblers and usually their families), while other people’s gambling wanders in and out of bounds set by norms regulating ‘appropriate’ behavior. This construction of problem and pathological gambling has been based on an established popular distinction between ‘problem drinking’ and ‘alcoholism’.

\textsuperscript{1} A caveat is that I don’t agree with Murphy that successful psychiatric science requires a \textit{general} account of abnormal mental functioning – different disorders constitute disorders for different kinds of reasons. Thanks to Harold Kincaid for this point.
Typically for research motivated by social concerns, the earliest research on problem and pathological gambling concentrated on establishing quantitative magnitudes – mainly, prevalence rates and aggregate social costs. This research has avoided mapping the problem/pathological gambling distinction onto any hypothesized distinction ‘internal to’ the minds or brains of disordered gamblers. Researchers have instead relied on operationalizations referenced to social-behavioral consequences. Studies gather subject samples of suitable size and representativeness for estimating proportions of target populations that gamble never, ‘occasionally’, ‘regularly’, and ‘very frequently’, and proportions that gamble more than the gamblers or their families wish they did. The first three categories refer to social norms, and the fourth category is not presupposed to be a strict subset of the third. What makes prevalence estimation scientific, along with sound analysis, is attention to cross-study comparability: similar methods are applied recurrently in different populations. We have evidence that prevalence estimates track a phenomenon stable enough for accumulation of knowledge just in case we discover similar category proportions in various populations after controlling for hypothetically relevant environmental conditions (e.g. availability of gambling opportunities).

Prevalence estimation is characteristic of a ‘phenomena counting’ stage of science, preceding experimental and theoretical refinement. Policy, clinical and diagnostic practices reliant on the problem/pathological distinction have not been able to wait for the science to mature. This does not mean that practice has ignored ongoing science pending its maturity. Clinicians and policy makers generally suppose that responsible policy at a given time t should reflect whatever scientists have agreed upon by t.
Furthermore, since funding for research on pathologies is mainly motivated by clinical imperatives, distinctions used in diagnosis condition the formulation of hypotheses. Thus scientific and clinical conceptualizations mutually interact, but clinical practices and principles dominate the interaction.

Let us idealize ‘scientific’ isolation of a phenomenon as the identification of a causal regularity, relation or disposition that holds under a specified range of conditions. Although clinical communities stabilize syndrome concepts in ways that depart further from this ideal than mature scientific research programs, clinicians’ convergence on shared diagnostic criteria is not unsystematic. First, practitioners report symptoms they observe occurring together in their accidental patient samples. These reports are published in journals of patient observations, which are periodically reviewed by meta-analysts in search of co-morbidity patterns that significantly recur. Where such patterns can’t be explained as consequences of the interactions of already established syndromes, they become the basis for new diagnoses. Once there is approximate consensus on a diagnosis and on the best treatment given current knowledge and technology, this is published in a diagnostic manual. Manual entries do more than summarize observations. They also refine diagnostic practice by supplying standard tests for confirmation. This allows clinicians to refer patients along through treatment networks using diagnostic

---

2 This formulation is intended as neutral among philosophies of science; it might refer to Humean regularities, to nomological relations among ‘natural kinds’, to Cartwright’s (1992) ‘Aristotelian natures’, or to functional dependencies among restricted parametric ranges of variables in models with specified domains of application.
labels that are interpreted in roughly the same way by everyone who consults the patient’s file.

Based on this practice, *DSM-IV* operationalizes ‘pathological gambling’ as:

A chronic inability to refrain from gambling to an extent that causes serious disruption to core life aspects such as career, health and family. A person is diagnosed as a probable pathological gambler if they agree with five or more of the following statements:

1. You have often gambled longer than you had planned.
2. You have often gambled until your last dollar was gone.
3. Thoughts of gambling have caused you to lose sleep.
4. You have used your income or savings to gamble while letting bills go unpaid.
5. You have made repeated, unsuccessful attempts to stop gambling.
6. You have broken the law or considered breaking the law to finance your gambling.
7. You have borrowed money to finance your gambling.
8. You have felt depressed or suicidal because of your gambling losses.
9. You have been remorseful after gambling.
10. You have gambled to get money to meet your financial obligations.
As with other *DSM* entries, this is an operationalization in a precise sense, intended as the basis for constructing diagnostic screens for administration to reporting patients. The belief that most people with serious gambling disorders will agree with five or more of the statements above is not based on scientific research but on anecdotal clinical lore, and on a tradition brought over from more extensive psychiatric experience with patterns in alcohol dependence.

‘Problem gambling’ is not defined in *DSM-IV*. Nevertheless, the US Committee on the Social and Economic Impact of Pathological Gambling, composed mainly of scientists, operationalizes problem gambling as “gambling behavior that results in any harmful effects to the gambler, his or her family, significant others, friends, co-workers etc.” (National Research Council 1999, 21). Though the Committee endorsed the *DSM* operationalization of ‘pathological gambling’ and explicitly defined ‘disordered gambling’ as the union of pathological and problem gambling, its operationalization of the latter is unconnected to any screen. Taken literally, it is clinically unhelpful: losing $5 on a football bet or being late for lunch due to a queue at the betting window constitute ‘harmful effects’ of gambling, however trivial. Since ‘problem gambling’ is intended to denote a warning condition for vulnerability to pathological gambling, one might expect that operationalizations of the two ideas should have a common basis: an assessor should be able to use the same screen by which she identifies pathological gamblers to identify problem gamblers, applying a lower threshold for the latter. However, screens based on the *DSM* reflect no principled underlying scale robust under transformations to alternative threshold tests. Clinicians do not treat the statements in the *DSM*
operationalization as equally diagnostic. Agreement with statements (5), (6) or (10) is
taken to indicate a probable serious gambling problem, and likewise for (8) if ‘depressed’
is interpreted clinically. This cannot be said of the other statements barring special
interpretations. These asymmetries make it unclear what ‘lowering the threshold’ on
*DSM*-based screens might systematically mean.

Despite its unclarity, there are practical motivations for retaining the ‘problem gambling’
construct. There is consensus in the gambling industry, in the treatment community and
among regulators that large numbers of gamblers occasionally lose more money than they
judge they can comfortably afford, while a much smaller proportion find their lives and
welfare catastrophically impaired by relentless cravings to gamble. There is thus an
acknowledged public health interest in determining whether there is a scientific basis for
hypothesizing a qualitative ‘jump’ between typical consumers of gambling services with
imperfect self-control, and truly addicted gamblers who should be treated according to
precedents for Axis-I psychiatric disorders. Most casinos aim to identify and then exclude
addicted gamblers, but not those who merely occasionally gamble more than they
intended to. If there is *not* a qualitative jump between these types, and any frequency of
self-control lapses has predictive significance for probability of developing a psychiatric
disorder, then casino and lottery operators carry potentially dramatic levels of ethical
responsibility. On the other hand, if pathological gambling is a distinctive and relatively
*sui generis* condition, then operators might hope that their problem is no trickier in
principle than aiming to keep people with tuberculosis off jetliners.
Conceptual questions about the relationship between minor and severe gambling problems have therefore been framed as questions about the ontogeny of the latter given rates of the former. Prevalence studies have proven largely uninformative here. They have generated consensus that, for any population, adolescent prevalence is at least three times higher than among adults. However, alcohol and drug use studies show that the majority of adolescents who consume at rates which in an adult would be regarded as reliable indicators of addiction endogenously reduce these rates by their late twenties. Thus, though it appears that most addicted adults were impulsively consuming adolescents, most impulsively consuming adolescents don’t become addicted adults.

Treatment community lore holds that pathological gamblers often enjoyed statistically deviant win frequencies in their first gambling experiences, which, it is conjectured, leads to over-estimation of expected payoffs that is for some reason difficult to un-learn (Collins 2003). However, this idea has never been tested in a longitudinal treatment.

This unsatisfactory conceptual impasse is not simply a function of the recentness of gambling behavior science. It reflects principled epistemological problems with the reliance on clinical screens in research.

3. Clinical screens as research instruments

For conditions that can cause serious harm, the imperative to help suffering people has higher social priority than scientific knowledge. In screening to identify pathological gamblers, the treatment community prefers to err on the side of diagnosing risk in some
people who aren’t really in trouble, in order to avoid false negatives. Clinical screens are open to criticism if they do not build in this bias. This creates a dilemma for scientists choosing screens to select research samples. Use of clinical screens will over-estimate prevalence rates. If scientists instead develop customized research screens that aim to avoid bias, clinical and research samples will fail to match. This is problematic. Subject recruitment for studies of low-prevalence conditions, such as pathological gambling, is expensive. Practical considerations require frequent reliance on samples already assembled for treatment purposes, especially when an aspect of a research project is evaluation of a therapeutic approach. Ethical complications then arise because research cannot be concealed from its subjects. If we re-classify clinical populations to correct for screen bias, this sends patients confusing signals that may undermine their recovery. Thus most studies of pathological gambling have used samples that were sorted into treatment and control groups using clinical screens. This has had several important consequences.

First, it has created major uncertainties in prevalence estimation. One instrument, the South Oaks Gambling Screen (SOGS) (Lesieur and Blume 1987), had featured in 90% of published studies as of a survey by Dickerson and Baron (2000). The SOGS can be self-administered by respondents, which is cost-efficient in large samples because it avoids the need for reliance on qualified administrators of structured clinical interviews. However, its tendency to harvest false positives appears to be large even for clinical screens. Most textbooks and surveys cite ‘typical’ pathological gambling prevalence estimates in jurisdictions with legal casinos as 1% to 4% of adults. Ladouceur (personal correspondence), a leading pioneer of gambling research, anecdotally estimates the false
positive rate produced by SOGS-based estimation at at least 1:1, and recent systematic evidence supports this guess. New screens have been developed which, while still based on DSM criteria, better match the findings of structured clinical interviews than the SOGS. Using one of these instruments (Ferris and Wynne 2001), the British Gambling Prevalence Survey (Wardle et al 2007) produced an estimate of 0.5%. This corresponds closely to the result of the most rigorous prevalence study ever conducted, the Gambling and Co-occurring Disorders component of the US National Comorbidity Survey (Kessler 2007). In that study, full structured clinical interviews administered to a random sample of over 9,000 adults generated a prevalence estimate of 0.7%.

Second, in considering any inference of causal relationships from relative efficacies of therapies, one must know the average or marginal severity of pathology to which the screen used to recruit subjects was sensitive. Consider two hypothetical screens A and B. Suppose that screen A consistently produces samples in which the median subject agrees with 8 statements in the DSM-IV operationalization, while the median subject recruited by screen B agrees with 6 of them. One would then have reason to expect that research using screen B will more likely suggest that a given policy or therapy is effective than research on that same policy or therapy that recruits and pools subjects by means of screen A, because the A-screened samples will contain higher proportions of severe cases. On the other hand, as noted, not all the DSM-IV criteria are of equal diagnostic weight. Suppose, for example, that more of the B-screened subjects agree with statements (5) and (6).
Third, a major research question, indicated above, is whether there is a qualitative jump, revealed by some quantitative discontinuity on some measurable parameter, between subgroups of disordered gamblers. However, as we saw, prevalence work based on clinical screens is ill suited to shedding light on this. Blaszczynski and Nower (2002) argue that there are three different kinds of pathological gambler, that each kind has a different etiology, and that people suffering from each kind respond best to different interventions. Where diagnostic screens are concerned, this suggests the possibility of different scales with discontinuities at different measurement points of magnitudes of different variables. So long as we are relying on operationalizations of folk concepts instead of functional relationships represented in exact models, we should therefore ideally run every test we think is important on several groups of subjects recruited using different screens, and run regressions. But then absence of a common underlying model would block attempts to estimate the contributions of different variable values to different group memberships.

Failure to apply a scientific model to identification of pathological gamblers also makes it difficult to integrate study of the phenomenon with investigations of mental health more generally. Considerable activity has aimed at estimating the tendency of pathological gamblers to manifest Axis-I comorbidities (substance dependence, depression, schizophrenia, and anti-social personality disorder). Discovering stable comorbidity rates across populations of pathological gamblers, if there are any, would be helpful for a number of reasons. First, it would help predict the likelihood that a given person is at risk for pathological gambling. Second, it would be relevant to design of treatments and interventions, possibly explaining some patterns of success and failure: interventions that
work for pathological gamblers who lack certain comorbidities might fail for others who don’t (see, e.g., Ladd and Petry 2003). Third, efforts to explain stable comorbidity patterns might lead us toward the explanation of pathological gambling itself if pathological gambling and comorbidities are sometimes or often consequences of common factors. Comorbidity data are what mainly motivate Blaszczynski and Nower’s thesis concerning different kinds of pathological gamblers. They report that across studies with larger numbers of subjects, stable proportions of pathological gamblers show comorbidity with, respectively (i) nothing, (ii) depression and other mood disorders, and (iii) anti-social personality disorder. Subjects in these groups respond differentially to therapies. Blaszczynski and Nower then argue, more controversially, that the different comorbid factors causally contribute to pathological gambling in different ways.

Most recent gambling policy attends to scientific research on gambling out of interest in the extent to which availability of commercial gambling opportunities and regulated features of casino games causally impact on pathological gambling rates and severity. As the perspective of Blaszczynski and Nower reminds us, however, questions about causality are inherently bound up with whether pathological gambling is a sui generis disorder or a secondary expression of less specific problems. To the extent that we expect the latter, then gambling regulatory policy may be the wrong instrument for addressing the public health issues.

4. Neuroscience and molecular genetics to the rescue
If current prevalence studies of pathological gambling by themselves shed little light on causal relationships, how might we do better both for the sake of scientific knowledge and improved treatment and policy? Brute force methodology, running controlled behavioral experiments on random subject samples in which suspected causal factors are systematically manipulated, is impractical due to the apparent low-frequency prevalence – much lower than formerly thought – of pathological gambling. The other way of inferring causes is to gather data by reference to an explicit causal model and then test the model using a tailored regression technique. To my knowledge this econometric approach has not yet been attempted in any pathological gambling prevalence study. There are two reasons for this. First, as noted, most pre-neuroscientific investigations of pathological gambling proceed on the basis of no underlying model. Second, the abstractness of the ‘pathological gambling’ construct, along with its multifarious demonstrated comorbidities, gives rise to potentially vicious endogeneity problems for model estimation. It is likely that the dependent variable of ultimate interest – disposition to pathological gambling – is determined to different degrees by different observed independent variables and these in turn partly co-determine one another. This limitation forces most prevalence researchers (like most social psychologists) to be content with weak tests of significance (e.g T-tests) on correlations, unable to employ stronger econometric tests that aim to isolate causal structure.

These problems are not insurmountable; on the cutting edge of gambling research they are indeed being gradually surmounted. As in cognitive and behavioral science generally,
two disciplinary clusters are coming increasingly to the rescue: neuroscience and molecular genetics.

Recent combined neurochemical and neuroeconomic models of addiction, specifically including gambling addiction, are surveyed in Ross et al (2008). It has been known for some time that addiction is correlated with abnormal levels of the neurotransmitter dopamine in ventral striatum. The dopamine circuit from midbrain to striatum has been more recently identified as responsible for predicting rewards, valuing potential rewards against alternatives, and preparing motor activity to procure rewards. New neuroimaging technologies yield sufficiently discriminating information about comparative neurotransmitter activity in addicts and non-addicts to test models of processes by which this circuit can capture the regulation of a person’s or rat’s molar behavior. All of these models share some general features. Essentially, they represent the mesolimbic dopamine reward system as gaining control of the organism’s molar behavior by chemically attenuating feedback serotonin (and other) circuits from frontal and prefrontal cortex which normally bid for attention to, and for scheduling consumption of, longer-range sources of reward. Fear responses from emotional centers such as the amygdala are also suppressed. The reward system achieves this mutiny by exploiting the discovery that, through relentlessly searching for cues to the arrival of an addictive target, and then organizing consumption of that target, it can reliably produce floods of dopamine in striatum. This constitutes the reward the system is evolved to maximize, which simultaneously overwhelms the functioning of normally rival circuits.
People who report greatest behavioral disturbance with respect to gambling closely resemble drug addicts in having hyperactive striatal dopamine responses in the presence of cues for gambling (drugs), including their own fantasizing about gambling (drugs), and hypoactive opponent neurotransmitter responses. This explains why seriously afflicted pathological gamblers without comorbid substance abuse display behavioral tendencies familiar from drug addiction. It is hypothesized on this basis that some proportion of people diagnosed as pathological gamblers by the *DSM* criteria are neurochemically addicted to gambling. Given the tendency of *DSM*-based screens to identify false positives, it is thus probable that the screens tend to capture mixes of addicts and neurochemically normal frequent gamblers. In this way they fail to cut nature at scientific joints, *whether by this idea we allude to a realist model of psychiatric natural kinds, or only to an ambition to write down reduced-form models that permit isolation of asymmetrically co-dependent variables that will be robust under econometric tests*. In the interdisciplinary matrix of neuroeconomics, economists tend to emphasize empiricist virtues by pursuing the second aim. Neuroscientists tend to try to isolate real entities and mechanisms. This difference does not prevent them from jointly converging on a shared basis for dividing the normative concept of pathological gambling.

Discovering that at least a subset of what psychiatrists have called pathological gambling is a pathology in the balance of power among neurotransmitter circuits is directly relevant to progress in understanding the disorder’s pathogenesis. There was already limited evidence from twin studies for a heritable aspect of pathological gambling. Based on examination of 3359 twin pairs, Eisen et al (1998) conclude that inherited factors explain
between 35% and 54% of reports of five \textit{DSM} symptoms that could be estimated statistically, 56% of the report of three or more symptoms, and 62% of diagnoses. Potenza \textit{et al} (2005) find a matched genetic contribution with near-perfect overlap to pathological gambling and major depression, while Slutske \textit{et al} (2000) report common genetic vulnerability to pathological gambling and alcohol dependence in male subjects.

All of this is at best suggestive. Tracing the phenomenon of severe pathological gambling to the dopamine circuit, however, enables molecular geneticists to hone in on the possible biological explainers of these data. Ibáñez \textit{et al} (2003) review seven association studies, all but one conducted by their group. An important property of their work is its motivation to provide separating evidence between conceptions of pathological gambling as an obsessive-compulsive disorder – the category in which \textit{DSM} places pathological gambling but \textit{not} drug addiction – and the family of dopaminergic pathologies that clearly include the classic addictions. In light of the reward system mutiny model, as they say, “genes relevant to the function of serotonergic, dopaminergic and noradrenergic systems could be considered as candidate genes in pathological gambling” (16). They find a promoter polymorphism sequence for expression of MOA-A protein, which has been associated with control of neurotransmitters found in the reward system, to be significantly increased in male pathological gamblers compared to controls.

“Interestingly,” they comment, “although serotonin is a preferred substrate for MOA-A, MOA-A is expressed in the brain mainly in dopaminergic neurons, raising the question of whether these allele variants are more likely to result in changes in serotonergic or dopaminergic transmission” (18). This comports suggestively with the mutiny model,
according to which it is attenuation of serotonin circuits by dopamine activity that mediates suppressed frontal control in addicts. In female pathological gamblers Perez de Castro et al (1997) discovered the DRD4 7-repeat allele, coding for less efficient receptors, to be significantly more frequent than in controls. Finally, Ibáñez et al cite a report due to Comings et al (1996) of “significant association between the Taq-A1 allele of the D2 dopamine receptor gene in pathological gamblers compared to controls,” and note that the Taq-A1 allele “has also been found to be associated with other impulsive-addictive-compulsive behaviors, leading some researchers to propose a Reward Deficiency Syndrome as an underlying genetic foundation for these disorders” (Ibáñez et al 2003, 18).

5. Conclusion

The philosophical interpretation I suggest is best motivated by these developments is as follows. There is good reason, independent of issues that divide realists and empiricists, for reconceptualizing the phenomenon that DSM-IV operationalizes as pathological gambling as a specific manifestation of disruption in frontal-cortical control of the mesolimbic dopamine system’s influence on molar behavior. By this proposal I do not intend a baroque claim to the effect that the authors of DSM-IV always intended ‘pathological gambling’ to refer to whatever unknown ‘constitutive essence’ unites the stereotypical cases, which has since turned out to be dopamine floods in ventral striatum. To the extent that it makes sense in the first place to talk about ‘intentions’ of ‘authors’ of referential conventions that emerge from complex institutional politics and evolving
diagnostic practice over many years, these ‘intentions’ were highly unspecific, and should be inferred directly from the general function of the *DSM*. That function is to optimally facilitate and standardize clinical reference, diagnosis and treatment of cases of patients with common symptoms, common etiologies of disturbances, and common response modalities to a common set of related interventions, with greater weight given to those whose suffering is most severe and chronic. Then, I claim, the empirical evidence suggests that the largest proportion of those ‘intended’ (in this sense) to be diagnosed as pathological gamblers by the *DSM* operationalization, including almost all of those whose suffering is severe, chronic, and recalcitrant to low-intensity therapy, are afflicted with neuroadapted hypoactivity of serotonergic circuits that normally inhibit impulsive behavior, and dopaminergic reward circuits that have learned to obsessively pursue and attend to predictors of gambling opportunities. Because of this common condition, almost all PGs are candidates for a common neuropharmacological treatment, identification of which should be (as it now is) a priority in applied research. A priority in basic research on pathological gambling should be (as it now is) the refinement and generalization of relevant dopamine learning models, so as to improve both predictive power and the depth of our explanation and integration of pathological gambling in the wider context of the behavioral and brain sciences.

This reduction of a social-behavioral syndrome to a relationship among neurotransmitter systems is not what philosophers often call ‘seamless’. Some patients to whom *DSM-IV* ‘intended’ to class as pathological gamblers stand to be re-classified, though their social and behavioral problems remain as before. I believe that the case is typical in this respect,
at least of all sciences that are partly regulated in their development by practical applications.

References


