Value-Based Decision Making in Mental Illness: A Meta-Analysis

Dahlia Mukherjee and Joseph W. Kable

Clinical Psychological Science 2014 2: 767 originally published online 23 May 2014
DOI: 10.1177/2167702614531580

Theoretical/Methodological/Review Article

Clinical Psychological Science
2014, Vol. 2(6) 767–782

DOI: 10.1177/2167702614531580
cpx.sagepub.com

Value-Based Decision Making in Mental Illness: A Meta-Analysis

Dahlia Mukherjee and Joseph W. Kable
Department of Psychology, University of Pennsylvania

Abstract

In this study, we assessed value-based decision making in individuals diagnosed with mental illness. Two meta-analyses were conducted of studies that used the Iowa Gambling Task (IGT) to assess value-based decision making. In the first meta-analysis (63 studies, combined $N = 4{,}978$), we compared IGT performance in healthy populations and populations with mental illness. In the second meta-analysis (40 studies, combined $N = 1{,}813$), we examined raw IGT performance scores as a function of type of mental illness. The first meta-analysis demonstrated that individuals with mental illness performed significantly worse than did healthy control individuals. The second meta-analysis demonstrated no performance differences based on type of mental illness. These findings suggest that value-based decision making is a promising target for transdiagnostic analyses of processes that go awry in mental illness. A critical priority for future work, given that impairment in the IGT could arise from changes in several decision processes, will be to investigate the specific decision processes affected in different mental illnesses.

Keywords

decision making, mental illness, frontal lobe lesions, meta-analysis

Received 9/25/13; Revision accepted 3/24/14

Research in mental illness is beginning to shift emphasis away from disorder symptoms and toward basic processes that can go awry across several disorders (Sanislow et al., 2010). Mental illness is currently organized according to clinical syndromes, but critics of this approach point to notable heterogeneity within syndromes and comorbidity and overlapping features across syndromes (for a more detailed review, see Follette, 1996). As an alternative to the syndrome approach, psychologists are increasingly studying—and sometimes even treating—basic processes that cut across traditional mental-illness categories. This cross-cutting approach has begun to shape funding priorities at the National Institute of Mental Health, for example, through the recently introduced research domain criteria proposal (Insel et al., 2010).

Within a process framework, one category of processes that might be worth investigating is the set involved in value-based decision making. We use the term *value-based decision making* to denote the set of processes that is the object of study in behavioral and experimental economics and in the psychology of decision making (including consumer behavior). We use this term to distinguish value-based decision making both from the cognitive psychology of reasoning and judgment and from the psychophysical study of perceptual decisions.

Two kinds of evidence suggest that processes involved in value-based decision making might be affected by mental illness. First, the neural circuitry implicated in value-based decision making overlaps with that known to be impaired in mental illness. Studies on the neuroscience of value-based decision making have focused on the fundamental role of medial frontal and orbitofrontal cortex, striatum, amygdala, and the modulatory neurotransmitter systems that project to these regions (Bartra, McGuire, & Kable, 2013; Kable & Glimcher, 2007, 2009; Rangel & Hare, 2010). It is interesting that these very same regions also demonstrate neurochemical and functional disruption in different mental illnesses (e.g., Dom, Sabbe, Hulstijn, & Van Den Brink, 2005; Verdejo-García & Bechara, 2009; Verdejo-García, Pérez-García, & Bechara,

Corresponding Author:
Dahlia Mukherjee, Department of Psychology, University of Pennsylvania, 3720 Walnut St., Room C6, Philadelphia, PA 19104
E-mail: dahliam@sas.upenn.edu

2006), including obsessive-compulsive disorder (OCD; Cavedini et al., 2002; Lawrence et al., 2006); depression (Grecius et al., 2007; Mayberg, 2006; Murphy et al., 2001); alcohol, cocaine, and stimulant abuse (Bechara et al., 2001; Volkow & Fowler, 2000); pathological gambling (Brand et al., 2005); and personality disorders (Raine, Lencz, Bihrle, LaCasse, & Colletti, 2000). Second, there is some empirical evidence for differences in value-based decision making between individuals with mental illness and healthy control individuals, with examples in schizophrenia (Sevy et al., 2007), OCD (Tolin, Abramowitz, Brigidi, & Foa, 2003), substance dependence (Bechara et al., 2001; Bickel & Marsch, 2001), and depression (Clark, Chamberlain, & Sahakian, 2009).

This theoretical and empirical evidence has led some researchers to propose that studies of value-based decision making hold much promise for disentangling the fundamental processes that go awry in different forms of mental illness. Previous work on value-based decision making in healthy individuals provides exactly the right tools—both the key theoretical constructs and the laboratory tasks to measure them—needed for this effort. Montague, Dolan, Friston, and Dayan (2012) have even coined a name for this nascent field, "computational psychiatry," and reviewed some of the early studies employing this research approach (for further qualitative reviews, see also Cavedini, Gorini, & Bellini, 2006; Sevy et al, 2007).

Although there is reason for excitement about computational psychiatry, there also remains significant uncertainty about how fruitful this research strategy might prove to be. Most researchers have compared a single population of individuals with mental illness with healthy control participants. The studies to date often involve small sample sizes, and this lack of power has led to inconsistencies across studies, which means that it is unclear whether value-based decision making is indeed impaired in mental illness. Furthermore, given that in most studies, researchers have investigated only a single clinical group, and that different studies use different tasks, it is unclear whether some forms of mental illness, compared with other forms, might demonstrate more impairment in value-based decision making. We conducted quantitative meta-analyses to answer these two main questions, namely, (a) whether people with mental illness display significantly impaired value-based decision making relative to healthy individuals, and (b) whether there are any differences in value-based decision making across populations with different mental illnesses. The answers to these questions should inform the prospects for computational psychiatry—that is, the extent to which researchers should study value-based decision making in mental illness at all and, if so, for which mental illnesses these studies might prove most important.

To answer these questions, we focused our meta-analyses on the one task sensitive to value-based decision-making processes that has been used widely across all types of mental illness, the Iowa Gambling Task (IGT; see the Method section for a full description of the criteria and search procedure used to select the IGT). The IGT is a standardized instrument that assesses decision making in ambiguous situations (Bechara, Damasio, Damasio, & Anderson, 1994). In the IGT, individuals choose cards from among four decks (decks labeled A, B, C, and D). Choices from two of the four decks (C and D) result in moderate gains as well as moderate losses. Choices from the other two decks (A and B) result in much higher gains as well as much higher losses. Consistent choices from decks C and D result in a net gain, whereas consistent choices from decks A and B result in a net loss. Thus, decks C and D are considered "advantageous," and decks A and B are considered "disadvantageous." Performance is typically characterized by the number of choices of the advantageous decks minus the number of choices of the disadvantageous decks. Participants are unaware of these facts and must learn to maximize their monetary gain on the basis of the feedback they receive after each choice. Typical performance evolves during the course of the task, and most healthy participants make more choices from the advantageous decks by the end of the task.

As this description should make clear, the IGT taps into many different aspects of value-based decision making, including one's tolerance for risk and ambiguity (Ellsberg, 1961; Holt & Laury, 2002; Levy, Snell, Nelson, Rustichini, & Glimcher, 2010), the degree to which one weights losses versus gains (Kahneman & Tversky, 1979), and how well one learns on the basis of positive and negative feedback (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schönberg, Daw, Joel, & O'Doherty, 2007; Vaidya, Knutson, O'Leary, Block, & Magnotta, 2013). This limits the IGT's specificity in terms of the target decision process affected (at least when overall performance is analyzed as in the studies we reviewed, as opposed to studies that have used computational modeling to tease apart these various factors; e.g., Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Fridberg et al., 2010; Weller, Levin, & Bechara, 2010). However, the flip side of this lack of specificity is sensitivity to many different aspects of value-based decision making. This sensitivity is useful for our purpose, which is to assess whether value-based decision-making processes are impaired at all in mental illness and to compare the degree of any impairment broadly across different mental illnesses. The IGT can thus serve to screen for potential impairments in value-based decision making and, therefore, provides a sensible starting point for the first quantitative comparison of value-based decision making across mental illnesses.

We performed two meta-analyses across all studies that used the IGT to assess decision making in a population with mental illness. The first meta-analysis looked at effect sizes for comparisons between healthy individuals and those with mental illness. We were particularly interested in whether there was a significant effect across all mental illnesses and whether effect sizes reliably differed across disorders. A follow-up meta-analysis evaluated the raw scores from the IGT across different mental illnesses, rather than the differences in performance against matched healthy participants. This meta-analysis allowed a direct comparison of performance across disorders.

Method

Selection of the IGT

Given that our goal was to assess and compare value-based decision making across different forms of mental illness, we first searched for decision-making tasks that had been used widely enough for this purpose. An initial PsychINFO screen through February 2011 used *decision making* and specific disorders (e.g., *obsessive-compulsive disorder* or *OCD*) as descriptors. This broad search identified three value-based decision-making tasks that had been widely used in studies of mental illness: the IGT, the Delay Discounting Task, and the Balloon Analogue Risk Task. Three further searches were conducted on PsychINFO through February 2011. The first search used the following descriptors: *Delayed Discounting Task* or *Kirby Delay Discounting Measure* or *Temporal Discounting Task* or *Discounting Task* or *Probability Discounting Task*. The second search used *Balloon Analogue Risk Task* or *Balloon Analogue Risk Taking Task* as descriptors. The third search used *Iowa Gambling Task* or *IGT* as descriptors. All three searches were restricted to search the adult population (keyword = *adult*). The selection of the final task or tasks was made on the basis of two criteria—the task had been used to assess impairment across mental-illness categories (e.g., OCD, mood disorder, eating disorder, substance-dependence disorder, pathological gambling, schizophrenia, and personality disorder) and there were at least three independent studies in which impairment was assessed within each disorder category by using that task. We felt it was important to be able to compare across mental illnesses with the same exact task, given that different value-based decision-making tasks assess different constructs or combinations of constructs.

From these searches, we found only one task that met our criteria: the IGT. Of the 40 articles reviewed that used the Balloon Analogue Risk Task, only 3 fulfilled criteria for the present study. Of the 149 articles reviewed that used the Delay Discounting Task, only populations with substance abuse/dependence were prominently featured (20 articles). A preliminary review of 282 articles on studies that used the IGT revealed that the task would meet our criteria. Although we had hoped to be able to compare mental-health disorders on more than one task of value-based decision making, the IGT provides a sensible starting point for the first quantitative comparison, given that this task was used to initially identify both deficits in populations with brain lesions and activations in functional imaging studies that were later proven relevant to several other tasks that assess value-based decision making.

Search procedure

The broad clinical population under investigation prevented the use of specific and narrow search terms. Thus, the behavioral task of interest, the IGT, was used as a search term, and the available literature was assessed with respect to clinical pathology. Potentially relevant studies were identified via Google Scholar and PsychINFO searches through January 2012 using the descriptors *Iowa Gambling Task* or *IGT* and restricting the search to the adult population (keyword = *adult*).

The database searches were supplemented in several ways to ensure comprehensiveness. The reference lists of relevant reviews, chapters, and articles were manually searched for potentially eligible studies. The electronic library of the second author, who specializes in decision making, was also searched. From a provisional list of included studies, four researchers who frequently publish relevant studies were identified, and Google Scholar and PsychINFO searches were performed using these authors' names as search terms. Several steps were also taken to address publication bias. Unpublished dissertations were included in the PsychINFO search. The four authors who frequently publish relevant studies were also contacted and asked whether they had any unpublished studies pertinent to the research question. One researcher (Davis, 2011) provided unpublished data included in the study. Conference abstracts from the Society for Neuroeconomics (2005–2012), Society for Neuroscience (2000–2012), and Cognitive Neuroscience Society (2003–2012) were also searched. These abstract searches generated one additional study (Dolan, Bechara, & Nathan, 2007) that fit all inclusion criteria and was included in the final meta-analyses.

Studies had to meet the following inclusion criteria to be included in either meta-analysis:

(a) The study was published in English (to ensure proper coding).
(b) The sample consisted only of adult participants. This ensures comparable neurodevelopmental baselines across samples.

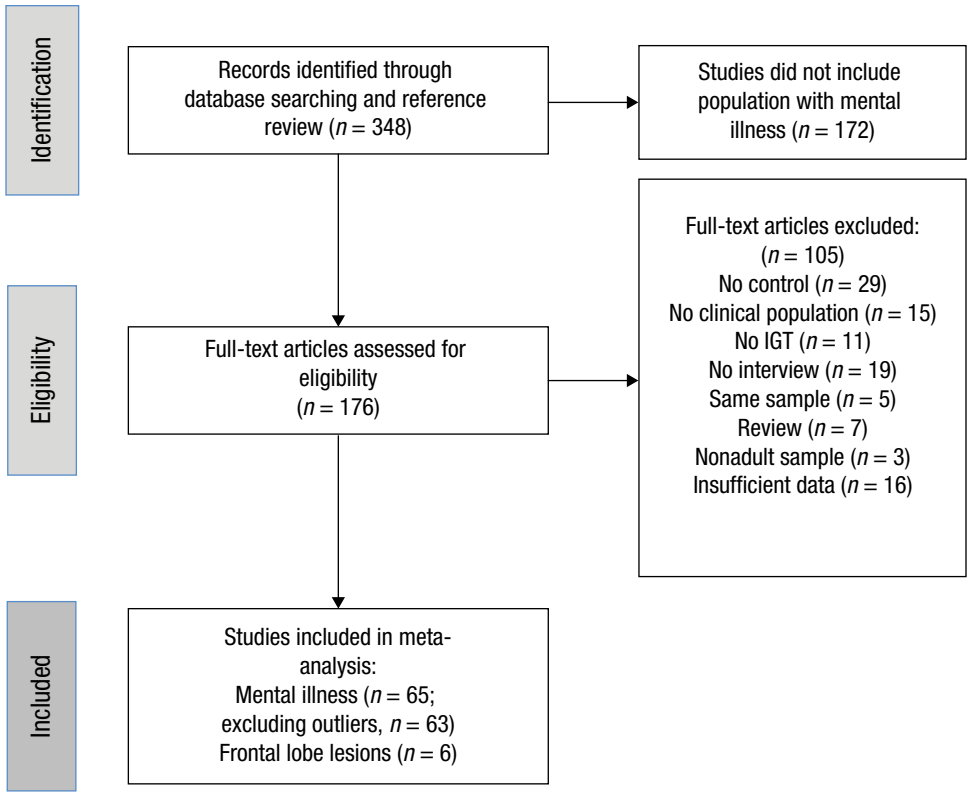

Fig. 1. Flowchart illustrating identification of included studies for the Iowa Gambling Task (IGT) between-group meta-analysis.

(c) For mental-illness diagnosis, only studies using clinical interview methods based on the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) guidelines were used. Self-report questionnaires could not be the sole source for establishing diagnosis. This helps improve reliability and validity of a diagnosis across studies.

For inclusion in the effect size meta-analysis, studies additionally had to include a comparison group of healthy participants and report sufficient details for an effect size to be calculated. For inclusion in the raw score meta-analysis, studies had to include one or both of the following: IGT net score means and standard deviations and block IGT scores and means for the clinical group under investigation (for a reference list of studies included in both meta-analyses, see Meta-Analyses References in the Supplemental Material available online).

Finally, to benchmark the IGT performance of populations with mental illness, we compared their performance with that of populations with brain lesions. Only studies using participants with ventromedial prefrontal cortex or frontal cortex lesions were selected; studies that involved traumatic brain injury were excluded from analysis. Given their known impairment on the IGT (Bechara, Damasio, Damasio, & Lee, 1999; Bechara et al., 1994; Bechara, Tranel, & Damasio 2000), this group was included as a comparison to assess the severity of any deficits in mental illnesses.

Selection of studies

The combined search for both meta-analyses yielded 348 studies, of which 176 were retained for evaluation for inclusion (see Figs. 1 and 2 for flowcharts illustrating identification of studies). Of these 176 studies, 6 studied populations with lesions and were used for comparison with populations with mental illness. Sixty studies were excluded from both meta-analyses for the following reasons: described a study that did not include a clinical population (n = 15), did not include the IGT as one of its decision-making assessment tasks (n = 11), did not use interview techniques and *DSM–IV* guidelines to determine mental-health diagnosis or relied on self-report measures (n = 19), was a repetition of the same sample described in another study (n = 5), was a review or theoretical article (n = 7), or included a population group other than adults (n = 3). For the effect size meta-analysis, 45 additional studies were excluded from analysis because

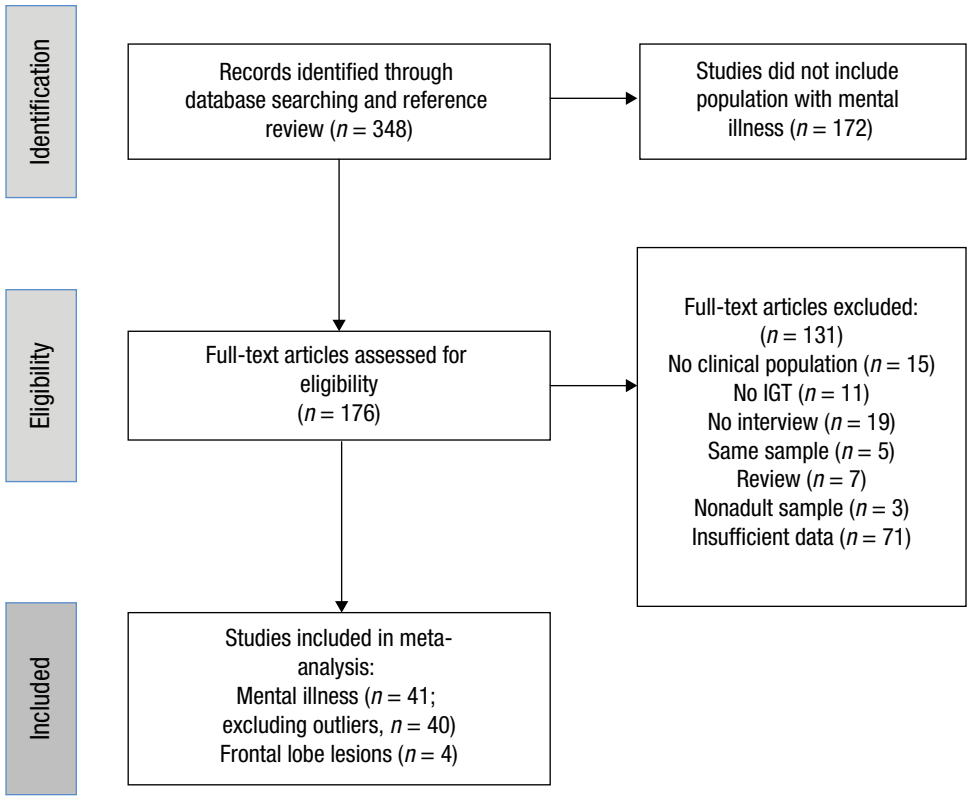

Fig. 2. Flowchart illustrating identification of included studies for the Iowa Gambling Task (IGT) mean performance meta-analysis.

they did not include a healthy adult comparison group (n = 29) or they reported inadequate data for the calculation of effect sizes (n = 16). For the raw score meta-analysis, 69 additional studies were excluded from analysis because they reported inadequate data for the calculation of performance scores. Thus, there were 65 articles, comprising 65 studies, that met criteria for the effect size meta-analysis and 41 articles, comprising 41 studies, that met criteria for the raw score meta-analysis. When a secondary source was available for a given study, the primary source was used to calculate the effect size unless reported data were insufficient. For comparison of effects sizes, 6 studies involving populations with lesions were available. For comparison of raw scores, only 4 of these studies were included because the others provided insufficient data to calculate net mean IGT performance and standard errors for the group with lesions. One author provided the necessary data for 3 of the 4 studies (Bechara et al., 1994; Bechara et al., 1999; Bechara et al., 2000).

Coding of studies

Given that one of our central questions was whether IGT performance varies with type of mental illness, studies were coded for both the specific type of primary mental-health disorder (OCDs, mood disorders, eating disorders, schizophrenia, substance abuse/dependence disorders, pathological gambling disorder, and personality disorders) and broadly into personality versus nonpersonality primary disorders. This broad distinction is interesting given the collapsing of Axis I and II disorders in the *DSM–5* (5th ed.; American Psychiatric Association, 2013). Lack of a difference between personality and nonpersonality disorders would be in alignment with the collapsing of Axis I and II categories in the *DSM–5*. Studies were also coded for other variables that could explain heterogeneity in effects and that were consistently reported in all studies. Because intelligence is known to affect value-based decision making (Burks, Carpenter, Goette, & Rustichini, 2009), each study was coded for whether an intelligence assessment was administered, and if intelligence was assessed, for whether there was a significant difference between the clinical and healthy groups. Studies were also coded for whether the study excluded participants with substance abuse/dependence and for whether the study excluded participants with traumatic brain injury or other neurological impairment (Dom et al., 2005). Comorbidity between mental disorders could be another important factor, as could whether individuals with mental-health disorders are currently

undergoing treatment and how long the individuals had been diagnosed with the disorder. Unfortunately, none of these factors were consistently reported across studies and, thus, could not be examined.

Independence of effect sizes

To meet the statistical assumption of independence of effect sizes, we took several steps to ensure that each study contributed only one effect size to each set of analyses (Lipsey & Wilson, 2001). For example, researchers may publish more than one article using the same data set or may include data used in a previous publication. Because this would violate the independence of effect sizes assumption, authors with multiple publications were contacted to provide information on whether completely independent clinical samples were recruited if they had investigated the same clinical group in more than one published article. Furthermore, in studies that included multiple clinical populations and a single healthy control group, only the clinical population with the largest sample size was used in the effect size meta-analysis, given that including multiple effect sizes calculated from the same control group would violate the independence assumption. For studies that subdivided their target clinical population into subgroups (e.g., alcohol dependence with personality disorder and alcohol dependence without personality disorder) and reported data for each subgroup, the combined means and standard deviations for those subgroups were calculated to increase sample size and ensure independence of effect sizes.

Meta-analytic procedure and analyses

Weighted mean effect sizes, heterogeneity analyses, and moderator analyses were conducted using Comprehensive Meta-Analysis, Version 2.2.046 (Borenstein, Hedges, Higgins, & Rothstein, 2005). Because the eligible studies used different samples, and methodologies, considerable heterogeneity of effects was expected. Fixed-effects models assume that the true effect size is the same in all studies, and any variability in effect sizes between studies is attributed to random error. By contrast, random-effects models assume that the true effect may vary systematically from study to study. Given the expected dispersion of effect sizes, random-effects analyses were used to model two aspects of the observed variance: random within-study variance and systematic between-study variance. Each effect size was weighted to account for its relative precision on the basis of the standard error of the effect size (within-study variance) and tau-squared (between-study variance).

Effect sizes. Effect sizes for between-group comparisons were coded such that a negative effect size indicated impaired decision-making performance in the clinical group relative to the control group. Hedges's *g* (Hedges, 1981) was employed as a measure of effect size. The conventions typically used to interpret Cohen's *d* can also be applied to Hedges's *g*: An effect size of 0.2 is considered small, 0.5 is considered moderate, and 0.8 is considered large (Cohen, 1988). For studies that did not provide sufficient data for an effect size calculation but reported nonsignificant results, an effect size of 0 was entered (Lipsey & Wilson, 2001). Given the range of study designs and purposes, methodological quality was not quantified or used in the weighting of effect sizes.

Raw scores. For the raw score meta-analysis, performance on the IGT was expressed as the number of selections of good decks minus the number of selections of bad decks. Mean scores were calculated. Mean performance can range widely, and healthy control participants typically score in the positive range.

Outliers. Final effect sizes or raw scores greater than or equal to 3 *SD* above or below the weighted mean were identified as outliers. Two outliers were detected (Dolan et al., 2007; Dom, De Wilde, Hulstijn & Sabbe, 2007) in the effect size meta-analysis. One outlier was detected for the raw score meta-analysis (Maurex et al., 2009). The results presented here exclude the outliers, such that the final effect size meta-analysis included 63 studies and the final raw score meta-analysis included 40 studies. The results did not differ significantly when the outliers were included. Details of all studies in the meta-analyses, including the three studies judged as outliers, are provided in Tables 1 and 2.

Publication bias. Publication bias (also called the "file drawer problem") presents a serious challenge for any meta-analysis. Studies with nonsignificant findings or small effect sizes have a decreased probability of being published, which can result in inflated estimates of effect size in meta-analyses of published findings. As detailed earlier, several steps were taken in the initial search stage to reduce the potential effect of publication bias (although, admittedly, these steps recovered only two unpublished studies).

We also tested statistically for the effects of publication bias in two ways. First, a funnel plot was created and visually examined. This graph plots the standard error for each study (determined by the study's sample size) against the study's effect size. The name "funnel plot" derives from the predicted presence of an inverted funnel. Studies with larger sample sizes provide more reliable estimates of the effect size and therefore should cluster more tightly around the mean toward the top of the plot, whereas smaller studies provide more variable

Table 1. Characteristics of Studies That Assessed Decision-Making Performance in Clinical Populations Using the Iowa Gambling Task

Study	Clinical type	Disorder type	Diagnosis	*N*	Intelligence assessment	Intelligence significant	Substance–abuse exclusion	TBI exclusion	Hedges's *g*
Adida et al., 2008	Psych	Mood	Bipolar	90	No	NA	No	Yes	–1.08
Adida et al., 2011	Psych	Mood	Bipolar	195	Yes	No	Yes	No	–0.57
Barry & Petry, 2008	Psych	Sub	Multiple	168	Yes	No	No	Yes	–0.49
Bechara et al., 1994	Lesion	Frontal	Frontal	50	No	NA	No	No	–1.50
Bechara et al., 2001	Lesion	Frontal	Frontal	45	Yes	No	Yes	No	–1.05
Boeka & Lokken, 2006	Psych	ED	BN	40	Yes	No	Yes	Yes	–0.92
Bolla et al., 2003	Psych	Sub	Cocaine	26	Yes	No	No	Yes	–0.32
Bolla et al., 2005	Psych	Sub	Marijuana	22	Yes	No	No	Yes	–1.16
Borges et al., 2011	Psych	Anx	OCD	118	Yes	Yes	No	No	–0.15
Brogan et al., 2010	Psych	ED	BN and AN	59	No	NA	No	No	–0.94
Cavedini et al., 2001	Psych	PG	PG	60	Yes	No	No	No	–1.24
Cavedini et al., 2002	Psych	Anx	OCD	68	No	NA	Yes	Yes	–0.95
Cavedini et al., 2004	Psych	ED	AN	141	No	NA	Yes	Yes	–0.94
Cavedini et al., 2010	Psych	Anx	OCD	66	No	NA	Yes	Yes	–1.44
Choi et al., 2011	Psych	Schiz	Schiz	48	Yes	No	Yes	Yes	–0.06
Clark et al., 2001	Psych	MD	Bipolar	45	Yes	Yes	Yes	Yes	–0.79
Clark et al., 2003	Lesion	Frontal	Frontal	62	Yes	No	No	No	–1.09
Da Rocha et al., 2011	Psych	Anx	OCD	214	Yes	No	Yes	No	–0.63
Davis, 2011	Psych	ED	BED	191	No	NA	No	No	–0.24
Dolan et al., 2007*	Psych	Sub	Multiple	68	No	NA	No	No	–3.15
Dom et al., 2007*	Psych	Sub	Alcohol	91	Yes	NR	Yes	Yes	–2.28
Easton et al., 2008	Psych	Sub	Alcohol	25	Yes	Yes	No	No	–1.53
Evans et al., 2005	Psych	Schiz	Schiz	38	Yes	NR	No	No	0.00
Forbush et al., 2008	Psych	PG	PG	59	Yes	Yes	No	No	0.00
Fridberg et al., 2010	Psych	Sub	Marijuana	32	Yes	No	No	Yes	–0.95
Gonzalez-Blanch et al., 2008	Psych	Schiz	Schiz	91	Yes	Yes	Yes	Yes	–0.04
Grant et al., 2000	Psych	Sub	Multiple	54	Yes	Yes	No	Yes	–0.60
Grisham et al., 2007	Psych	Anx	OCD	60	Yes	Yes	Yes	No	–0.15
Guillaume et al., 2010	Psych	ED	BN and AN	170	Yes	Yes	No	No	0.00
Haaland & Landro, 2007	Psych	PD	BPD	35	Yes	Yes	No	Yes	–1.49
Hanson et al., 2008	Psych	Sub	Multiple	81	Yes	No	Yes	Yes	–0.54
Jollant et al., 2005	Psych	MD	Bipolar	107	Yes	No	No	Yes	–0.29
Kertzman et al., 2011	Psych	PG	PG	108	No	NA	Yes	No	–0.65
Kjome et al., 2010	Psych	Sub	Cocaine	86	Yes	Yes	No	No	–0.81
Lane et al., 2010	Psych	Sub	Cocaine	33	No	No	No	No	–0.70
Lawrence et al., 2006	Psych	Anx	OCD	79	Yes	No	Yes	Yes	–0.02
Liao et al., 2009	Psych	ED	BN	77	Yes	No	No	Yes	–0.50
Linnet et al., 2006	Psych	PG	PG	100	No	NA	No	No	–0.34
Linnet et al., 2011	Psych	PG	PG	30	No	NA	Yes	No	0.00
Loeber et al., 2009	Psych	Sub	Alcohol	84	Yes	No	No	Yes	–0.02
MacPherson et al., 2009	Lesion	Frontal	Frontal	38	No	NA	No	No	–0.89
Malloy-Diniz et al., 2009	Psych	MD	Bipolar	89	Yes	No	No	No	–1.09
Malloy-Diniz et al., 2011	Psych	MD	Bipolar	189	Yes	No	No	No	–0.69
Manes et al., 2002	Lesion	Frontal	Frontal	32	Yes	No	Yes	No	–1.28
Martino et al., 2007	Psych	Schiz	Schiz	36	Yes	No	Yes	Yes	–0.60
Martino et al., 2011	Psych	MD	Bipolar	119	Yes	No	Yes	Yes	–0.04
Maurex et al., 2009	Psych	PD	BPD	78	No	NA	Yes	No	–0.37
Mazas et al., 2000	Psych	PD	ASPD	53	Yes	No	No	No	–1.10
Miranda et al., 2009	Psych	Sub	Alcohol	60	Yes	Yes	No	No	–0.72
Must et al., 2006	Psych	MD	MDD	50	No	NA	Yes	No	–1.35

(continued)

Table 1. (Continued)

Study	Clinical type	Disorder type	Diagnosis	*N*	Intelligence assessment	Intelligence significant	Substance–abuse exclusion	TBI exclusion	Hedges's *g*
Nakamura et al., 2008	Psych	Schiz	Schiz	49	Yes	Yes	No	No	–0.76
Nielen et al., 2002	Psych	Anx	OCD	53	Yes	No	No	No	0.00
Petry et al., 1998	Psych	Sub	Heroin	93	Yes	No	No	No	–0.38
Pirastu et al., 2006	Psych	Sub	Opiate	69	Yes	Yes	No	Yes	–0.67
Premkumar et al., 2008	Psych	Schiz	Schiz	100	Yes	Yes	Yes	Yes	–0.48
Premkumar et al., 2010	Psych	Schiz	Schiz	45	Yes	NR	Yes	Yes	–0.47
Raffard et al., 2011	Psych	Schiz	Schiz	128	Yes	Yes	No	No	–0.46
Ritter et al., 2004	Psych	Schiz	Schiz	35	Yes	No	No	No	–0.74
Rodríguez-Sánchez et al., 2005	Psych	Schiz	Schiz	102	Yes	No	Yes	Yes	–0.13
Salgado et al., 2009	Psych	Sub	Alcohol	61	Yes	No	No	Yes	–0.94
Sevy et al., 2007	Psych	Schiz	Schiz	47	Yes	Yes	No	Yes	–0.35
Shirayama et al., 2010	Psych	Schiz	Schiz	37	Yes	No	Yes	Yes	–0.43
Shurman et al., 2005	Psych	Schiz	Schiz	49	No	NA	Yes	Yes	–1.54
Starcke et al., 2010	Psych	Anx	OCD	45	Yes	No	Yes	Yes	–1.07
Vadhan et al., 2009	Psych	Sub	Cocaine	46	Yes	Yes	No	No	–1.09
Van Toor et al., 2011	Psych	Sub	Multiple	62	No	NA	No	No	–0.94
Wesley et al., 2011	Psych	Sub	Marijuana	32	Yes	No	No	Yes	–0.28
Woicik et al., 2009	Psych	Sub	Cocaine	90	Yes	No	No	Yes	–0.08
Xi et al., 2011	Lesion	Frontal	Frontal	46	Yes	No	Yes	No	–0.92
Yip et al., 2009	Psych	Schiz	Schiz	63	Yes	Yes	Yes	No	–0.70
Zhang et al., 2011	Psych	Sub	Heroin	39	Yes	No	No	No	–0.64

Note: AN = anorexia nervosa; Anx = anxiety disorder; ASPD = antisocial personality disorder; BED = binge eating disorder; BN = bulimia nervosa; BPD = borderline personality disorder; ED = eating disorder; Frontal = frontal lobe lesion; Lesion = frontal lobe lesions or ventro medial lesions; MD = mood disorder; MDD = major depressive disorder; Multiple = multiple substance use; NA = not applicable; NR = not reported; OCD = obsessive-compulsive disorder; PD = personality disorder; PG = pathological gambling; Psych = mental illness; Schiz = schizophrenia; Sub = substance abuse/dependence; TBI = traumatic brain injury. Asterisks indicate studies excluded as outliers. For a reference list of studies included in the meta-analyses, see Meta-Analyses References in the Supplemental Material.

estimates and therefore should scatter more widely around the mean toward the bottom of the plot. In the presence of publication bias, the plot becomes asymmetrical, typically with fewer small-sample-sized studies than would be predicted with effect sizes smaller than the mean. The trim-and-fill procedure was then applied to the funnel plot (Duval & Tweedie, 2000). This procedure calculates the likely number of missing studies on the basis of the asymmetry in the funnel plot and produces an effect size and confidence interval adjusted to account for these missing studies. An important caveat to the use of these procedures is that funnel plots and the trim-and-fill procedure assume homogeneity of effect sizes. Heterogeneous data sets violate this assumption; thus, the use of these techniques in such cases (which include the present cases) should be interpreted with caution.

Publication bias was also examined using classic fail-safe values (Rosenthal, 1979). The fail-safe value determines the number of missing studies with a mean effect of 0 that would need to be added to the analysis before the two-tailed p value would be greater than .05. Tolerance levels were also calculated on the basis of the equation $5K + 10$, where K is the number of observed studies, proposed by Rosenthal (1979) to determine what would be considered an unlikely number of nonsignificant studies.

Homogeneity of effect sizes. The present data set was tested for homogeneity of effect sizes using the Q statistic (Hedges & Olkin, 1985) and the I^2 statistic (Cooper, 2010; Higgins & Thompson, 2002; Lipsey & Wilson, 2001). The Q statistic has a chi-square distribution and tests whether the observed dispersion is significantly larger than the expected dispersion based on within-study error. A significant Q statistic suggests that the distribution of effect sizes around the mean is greater than what would be predicted from sampling error alone. The I^2 statistic estimates the percentage of the variance that is attributable to between-studies variability as opposed to within-studies sampling error; percentages of 25, 50, and 75 generally indicate low, moderate, and high degrees of heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003).

Table 2. Characteristics of Studies That Reported Raw Mean Performance on the Iowa Gambling Task (IGT) in Clinical Populations

Study name	Disorder type	Diagnosis	*N*	Mean IGT performance
Adida et al., 2008	Mood	Bipolar	45	–1.50 (3.94)
Alfonso et al., 2011	Substance	Polysubstance	34	–6.72 (2.49)
Alvarez-Moya et al., 2011	Gambling	Pathological gambling	88	–1.10 (2.75)
Barry & Petry, 2008	Substance	Polysubstance	131	2.30 (1.76)
Bechara et al., 1994	Lesion	Frontal	6	–25.5 (11.54)
Bechara et al., 1999	Lesion	Frontal	19	–20.67 (6.76)
Bechara et al., 2000	Lesion	Frontal	10	–10.60 (2.4)
Bolla et al., 2003	Substance	Cocaine	13	6.17 (7.13)
Bolla et al., 2005	Substance	Marijuana	11	8.47 (4.37)
Borges et al., 2011	Anxiety	OCD	101	–3.71 (2.08)
Clark et al., 2003	Lesion	Frontal	41	–1.07 (3.26)
da Rocha et al., 2008	Anxiety	OCD	49	–2.29 (1.77)
da Rocha et al., 2011	Anxiety	OCD	107	–4.96 (1.24)
Davis, 2011	Eating disorder	Binge eating disorder	85	3.92 (2.85)
Dolan et al., 2007	Substance	Polysubstance	38	–2.10 (1.26)
Dom et al., 2007	Substance	Alcohol	38	2.40 (0.83)
Gonzalez-Blanch et al., 2008	Psychotic disorder	Schizophrenia	70	–1.10 (3.12)
Grant et al., 2000	Substance	Polysubstance	30	10.20 (4.70)
Grisham et al., 2007	Anxiety	OCD	30	5.28 (1.20)
Haaland & Landro, 2007	Personality disorder	Borderline personality disorder	20	–9.85 (5.44)
Jollant et al., 2005	Mood	Bipolar	25	9.20 (5.12)
Kjome et al., 2010	Substance	Cocaine	66	0.09 (2.77)
Linnet et al., 2006	Gambling	Pathological gambling	61	–0.31 (3.13)
Loeber et al., 2009	Substance	Alcohol	48	0.90 (1.16)
Malloy-Diniz et al., 2009	Mood	Bipolar	36	–1.03 (4.34)
Malloy-Diniz et al., 2011	Mood	Bipolar	95	3.89 (2.49)
Martino et al., 2007	Psychotic disorder	Schizophrenia	21	0.76 (6.12)
Maurex et al., 2009*	Personality disorder	Borderline personality disorder	48	18.90 (4.03)
Mazas et al., 2000	Personality disorder	Antisocial personality disorder	21	2.95 (3.79)
McNeely et al., 2008	Mood	Major depressive disorder	6	–3.10 (1.22)
Miranda et al., 2009	Substance	Alcohol	39	1.43 (3.28)
Nakamura et al., 2008	Psychotic disorder	Schizophrenia	24	–3.83 (5.54)
Pirastu et al., 2006	Substance	Opioid	48	11.30 (0.91)
Premkumar et al., 2008	Psychotic disorder	Schizophrenia	75	4.45 (1.39)
Premkumar et al., 2010	Psychotic disorder	Schizophrenia	30	2.80 (2.16)
Ritter et al., 2004	Psychotic disorder	Schizophrenia	20	–5.20 (4.41)
Rodriguez-Sanchez et al., 2005	Psychotic disorder	Schizophrenia	80	–1.63 (3.09)
Salgado et al., 2009	Substance	Alcohol	31	1.03 (2.62)
Sevy et al., 2007	Psychotic disorder	Schizophrenia	27	–5.00 (3.46)
Shirayama et al., 2010	Psychotic disorder	Schizophrenia	19	–4.74 (3.10)
Shurman et al., 2005	Psychotic disorder	Schizophrenia	39	1.90 (3.01)
Starcke et al., 2010	Anxiety	OCD	23	–1.50 (5.62)
van Toor et al., 2011	Substance	Polysubstance	31	–3.46 (4.73)
Wesley et al., 2011	Substance	Marijuana	16	–3.38 (2.18)
Yip et al., 2009	Psychotic disorder	Schizophrenia	42	5.20 (4.17)

Note: Standard errors shown in parentheses. OCD = obsessive-compulsive disorder. Asterisk indicates study excluded as an outlier. For a reference list of studies included in the meta-analyses, see Meta-Analyses References in the Supplemental Material.

Moderator analyses. Given evidence of substantial heterogeneity of effects sizes, moderator analyses were conducted on variables that might be associated with study effects and were consistently reported across studies (see the Coding of Studies section). Analysis of variance was conducted for categorical moderators using a mixed-effects model for each variable hypothesized to influence the effect size. This model consisted of a random-effects model, which combined studies within each subgroup, and a fixed-effect model that combined subgroups to determine the overall effect. Where applicable, the strength of differences based on moderator analyses was calculated using Cohen's *d*.

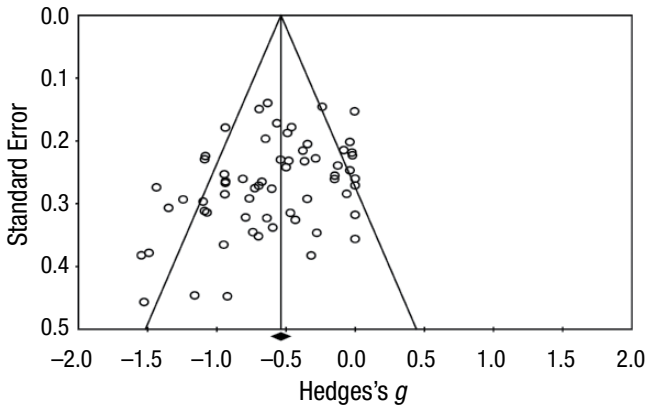

Fig. 3. Effect size meta-analysis funnel plot.

Results

Sixty-three studies contributed to the effect size meta-analysis and 40 studies to the raw score meta-analysis. Study and sample characteristics are presented in Tables 1 and 2, respectively.

Results for effect size meta-analysis

The results of the random-effects model for effect size indicated that participants with mental illness performed reliably worse on the IGT than did healthy control participants, with a moderately large effect size. Individual study effect sizes ranged from 0 to –1.55 (negative effect sizes indicate impaired performance in the clinical population). The average effect size was –0.58, 95% confidence interval (CI) = [–0.68, –0.48], $p < .001$. Cohen's U3 provides an intuitive metric to comprehend the magnitude of this effect size. A magnitude of –0.58 implies that 73% of participants in the clinical population could be expected to perform worse on the IGT than the mean performance level of healthy control participants (Lipsey & Wilson, 2001).

Although there was evidence of publication bias, the difference between individuals with mental illness and healthy control individuals was robust to this bias. The fail-safe value was 5,172, far exceeding the proposed tolerance levels of what would be considered an unlikely number of nonsignificant studies (350). The funnel plot was asymmetric, with an absence of potential studies on the lower right-hand side of the funnel (see Fig. 3); trim-and-fill procedures suggested that 13 studies with effect sizes to the right of the mean (more strongly positive) were missing. The corrected average effect size based on the trim-and-fill procedure (Duvall & Tweedie, 2000) was –0.44, 95% CI = [–0.55, –0.33].

We expected heterogeneous effect sizes, given that the populations comprised diverse mental illnesses. The *Q* statistic indicated significant heterogeneity among the effect sizes ($p < .001$). The I^2 value indicated moderate levels of heterogeneity; 60.71% of the variance in effect sizes was attributable to between-study variance.

Because the *Q* statistic and I^2 value indicated significant heterogeneity, an analysis of potential moderators was conducted to assess whether effect sizes differed on the basis of study characteristics (see Table 3). Neither type of mental illness, $Q(6) = 4.60$, $p = .60$, nor personality/other type, $Q(1) = 1.79$, $p = .18$, $d = 0.34$, was a significant moderator. However, in the case of personality/other type, lack of power may be a potential reason for the lack of significant findings. The trend was for the effect size in populations with nonpersonality disorders, formerly Axis I disorders ($g = –0.56$, $n = 60$), to be lower than that in populations with personality disorders, formerly Axis II disorders ($g = –0.90$, $n = 3$).

A significant moderator of effect size was the assessment of intellectual functioning. Studies that did not assess intellectual functioning reported significantly more impaired decision-making performance ($g = –0.86$, $n = 13$) than did studies that assessed intellectual functioning ($g = –0.51$, $n = 50$), $Q(1) = 8.58$, $p = .003$, $d = 0.35$. Among the 50 studies that assessed intellectual functioning, 18 reported a significant difference in intellectual performance between individuals with mental illness and healthy control individuals. These 18 studies also used intelligence as a covariate in data analysis, which may explain why there was no difference in effect size between studies that reported a significant difference in intellectual functioning ($g = –0.53$, $n = 18$) and studies that reported no difference ($g = –0.50$, $n = 32$), $Q(1) = 0.12$, $p = .72$, $d = 0.03$.

Neither substance-use exclusion, $Q(1) = 0.02$, $p = .90$, $d = 0.02$, nor exclusion for traumatic brain injury, $Q(1) = 0.005$, $p = .94$, $d = 0.02$, was found to be a significant moderator of effect sizes.

Finally, we compared the performance of individuals with mental illness with those with frontal lobe lesions.

Table 3. Analyses of Moderation for Effect Size Meta-analysis

Moderator	Study (*n*)	Hedges's *g*	95% confidence interval	*Q* (*df*)	*p*
Diagnosis	63			4.6 (6)	.60
Obsessive-compulsive disorder	8	–0.54**	[–0.81, –0.26]		
Eating disorder	6	–0.53**	[–0.84, –0.21]		
Mood disorder	8	–0.70***	[–0.98, –0.44]		
Pathological gambling disorder	5	–0.46	[–0.82, 0.10]		
Personality disorder	3	–0.90**	[–1.39, –0.41]		
Substance-dependence disorder	19	–0.63***	[–0.82, –0.44]		
Schizophrenia	14	–0.45***	[–0.68, –0.23]		
Nonpersonality vs. personality	63			1.79 (1)	.18
Nonpersonality disorder	60	–0.56***	[–0.66, –0.46]		
Personality disorder	3	–0.90***	[–1.38, –0.42]		
Intelligence assessment	63			8.58 (1)	.003**
No	13	–0.86***	[–1.07, –0.65]		
Yes	50	–0.51***	[–0.61, –0.40]		
Intelligence significant	50			2.4 (1)	.35
Yes	18	–0.53***	[–0.71, –0.35]		
No	30	–0.52***	[–0.66, –0.39]		
Not reported	2	–0.15	[–0.64, 0.33]		
Substance-use exclusion	63			1.08 (1)	.58
Yes	25	–0.58***	[–0.74, –0.41]		
No	38	–0.60***	[–0.74, –0.47]		
TBI/neuropsychological deficit exclusion	63			1.06 (1)	.60
Yes	32	–0.58***	[–0.73, –0.43]		
No	31	–0.60***	[–0.75, –0.46]		
Administration of IGT	63			1.40 (1)	.50
Computer	51	–0.56***	[–0.67, –0.44]		
Hand	12	–0.72***	[–1.00, –0.45]		
Clinical population	69			6.57 (1)	.01**
Lesion	6	–1.10***	[–1.48, –0.72]		
Mental illness	63	–0.58***	[–0.67, –0.48]		

Note: TBI = traumatic brain injury.
**$p < .01$. ***$p < .001$.

We combined the 63 mental-illness studies with 6 studies involving groups with frontal lesions and conducted a moderator analysis with type of clinical population as a moderator (lesion vs. mental illness). Type of clinical population proved to be a significant moderator, $Q(1) = 6.57$, $p = .01$, $d = 0.52$; the population with lesions performed significantly worse than did the population with mental illness. Thus, although the individuals with mental illness performed significantly worse than did the healthy control individuals, their deficits were not as large as those in individuals with frontal lobe lesions (see Fig. 4 for comparison of task performance among the clinical groups—nonpersonality disorder, personality disorder, and lesion).

Results for raw score meta-analysis

Performance on the IGT was quantified as the number of "good" deck choices (C + D) minus the number of "bad" deck choices (A + B). For 100 trials, this score can therefore vary from –100 to 100. Despite high variability in net IGT scores across populations, the mean net IGT score for healthy control individuals is usually a high net positive gain, on the order of 20 (Bechara et al., 1994), with higher scores implying better value-based decision-making strategies. For the population with mental illness, mean performances in individual studies ranged from –6.72 to 10.20. The average performance across all 40 studies was 0.45 ($SE = 0.88$).

Although there was evidence of publication bias (see Fig. 5), this did not seem to have a large effect on mean performance estimates. Trim-and-fill procedures suggested that 15 studies with raw scores to the right of mean (more strongly positive) were missing. The corrected average performance was 3.74, 95% CI = [–1.93, –5.54].

There was evidence for heterogeneity in effects across studies. The *Q* statistic indicated significant heterogeneity

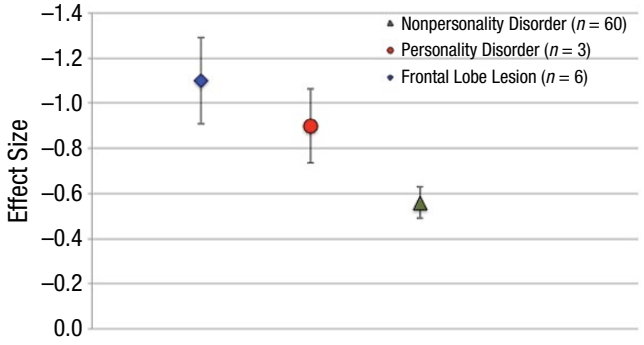

Fig. 4. Performance on Iowa Gambling Task as a function of clinical group. Negative values indicate impairment. Error bars represent standard errors.

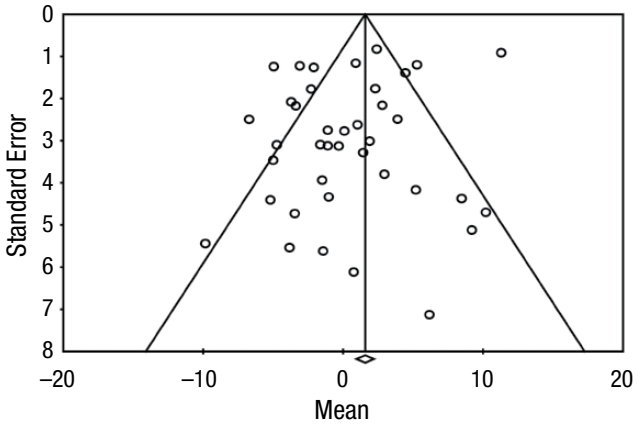

Fig. 5. Raw score meta-analysis funnel plot of standard error by mean.

($p < .001$). The I^2 value indicated high levels of heterogeneity, with 83.70% of the variance in raw scores attributable to between-study variance. Because both statistics indicated heterogeneity, we again conducted analyses of potential moderators (see Table 4). However, type of mental illness ($n = 40$) did not explain the heterogeneity in the IGT mean score, $Q(6) = 2.32$, $p = .90$. Personality disorders versus nonpersonality disorders did not moderate the effect either, $Q(1) = 0.38$, $p = .54$, $d = 0.84$. The direction of the difference was similar to the trend observed in the effect size meta-analysis; populations with nonpersonality disorders ($g = 0.57$, $n = 38$) performed better on the IGT than did populations with personality disorders ($g = –2.33$, $n = 2$).

Finally, we compared the performance of individuals with mental illness with those with frontal lobe lesions. We combined the 40 studies of mental illness with 4 studies involving groups with lesions and conducted a moderator analysis with type of clinical population as a moderator (lesion vs. mental illness). Type of clinical population proved to be a significant moderator, $Q(1) = 23.35$, $p < .001$, $d = 6.62$; the population with lesions again performed significantly worse than did the population with mental illness.

Discussion

This quantitative review answers two broad questions. The first concerns whether individuals with mental illness demonstrate impaired value-based decision making, as assessed by the IGT, relative to healthy individuals. The effect size meta-analysis demonstrated that performance on the IGT was significantly impaired in individuals with mental illness, although the effect size was moderate and the deficit was not as severe as in individuals with frontal lesions. The second question concerns whether, within the group with mental illness, the severity of impairment differs across types of mental illness. Given that different matched comparison groups might be used for different mental illnesses, the second meta-analysis, which directly compared the raw scores on the IGT in different mental illnesses, provides the clearest answer to this question. The raw score meta-analysis surprisingly did not demonstrate any significant differences based on type of mental illness.

The finding that value-based decision making is significantly impaired in mental illness may not come as a surprise to most readers. Qualitative reviews of decision-making behavior in specific disorders, such as schizophrenia (Sevy et al., 2007), OCD (Chamberlain et al., 2007), and substance-use disorder (Dom et al., 2005), all align with the present findings. However, qualitative reviews can fall prey to selection or publication bias, factors that careful, quantitative meta-analyses can address. To our knowledge, the present study is the first quantitative meta-analysis to verify that value-based decision making, as measured by any task, is consistently impaired in mental illness.

The present study is also novel in that it compared value-based decision-making performance across different mental illnesses. At present, only the IGT has been tested in a wide enough range of disorders to permit such a comparison. Here, our findings are perhaps more surprising: We did not find strong evidence for differential impairment on the IGT in different mental illnesses. Diagnosis was not a significant moderator in either the effect size or the raw score meta-analyses. Before we return to potential explanations for this lack of differential impairment, we first discuss two potential moderators that we did observe.

There was significant heterogeneity in size of IGT impairment across studies. Despite this heterogeneity, however, neither diagnosis nor personality versus nonpersonality type moderated IGT performance in the group with mental illness. There was, however, a trend in

Table 4. Analyses of Moderation for Raw Score Meta-analysis

Moderator	Study (*n*)	Effect size	95% confidence interval	*Q* (*df*)	*p*
Disorder	40			2.32 (6)	.9
Obsessive-compulsive disorder	5	–1.35	[–0.97, –0.09]		
Eating disorder	1	3.92	[–0.70, 14.86]		
Mood disorder	5	0.9	[–4.25, 6.04]		
Pathological gambling	2	–0.72	[–1.18, 0.14]		
Personality disorder	2	–2.37	[–11.54, 6.80]		
Substance dependence	14	1.72	[–1.2, 4.64]		
Schizophrenia	11	–0.28	[–3.8, 3.24]		
Nonpersonality vs. personality	40			0.38 (1)	.54
Nonpersonality disorder	38	0.57	[–1.19, 2.32]		
Personality disorder	2	–2.33	[–11.34, 6.67]		
Clinical group	44			23.35 (1)	< .001
Lesion	4	–14.12	[–19.74, –8.49]		
Mental illness	40	0.45	[–1.35, 2.26]		

both the effect size meta-analysis and the raw score meta-analysis for people with personality disorders to be more impaired on the IGT relative to people with other disorders. These findings should be interpreted with caution, given that only three studies of personality disorders were included in the first meta-analysis and that only two studies were included in the second. However, future studies focusing on value-based decision making in personality disorders are warranted, given the trend in the current findings. The current findings point to the possibility that people with personality disorders may experience more severe impairment than do those with other disorders. This would be consistent with the fact that personality disorders are more chronic and treatment resistant, whereas nonpersonality disorders usually have a more sudden onset and less prolonged time course. In fact, the IGT impairment in the small number of personality disorder studies in our meta-analyses was almost as large as that observed in patients with frontal lobe lesions, and individual studies of borderline personality disorder and antisocial personality disorder have uncovered large effects on learning from rewards or punishments (Rilling, King-Casas, & Sanfey, 2008). Given the comorbidity between personality disorders and nonpersonality disorders, our findings also demonstrate the importance of assessing and reporting comorbidity in future work. Indeed, it is possible that the IGT impairments observed in people with nonpersonality disorders in the current meta-analyses are due in part to comorbid personality disorders.

The second potential moderator, which was significant in the effect size meta-analysis, was assessment of general intellectual functioning. Studies assessing intellectual functioning reported smaller levels of IGT impairment compared with studies that did not assess intellectual functioning. Within the studies that did assess intellectual functioning, there was no difference in the size of IGT impairment between studies that observed a significant difference in intellectual functioning and studies that did not find such a difference. This is likely because in those studies that did establish a significant difference, intellectual functioning was used as a covariate while assessing IGT performance. Thus, when intellectual functioning is not assessed and controlled for, the degree of decision-making impairment on the IGT appears to be inflated. These findings suggest that the IGT is sensitive to not only value-based decision-making processes but also general intellectual abilities. In fact, several researchers have previously suggested that deficits on the IGT might indicate deficits in basic cognitive abilities, such as working memory (Hinson, Jameson, & Whitney, 2002; Jameson, Hinson, & Whitney, 2004). These findings illustrate the importance of assessing and controlling for intellectual functioning in studies of value-based decision making. It is critical to note, however, that even when assessing and controlling for intellectual functioning, impairment on the IGT is observed in mental illness.

Returning to our central finding, why might there be widespread impairments on the IGT across all mental illnesses, with no significant evidence for differential impairments? Although this null result (i.e., lack of differential impairment) should be interpreted with caution, a plausible interpretation arises from recognizing the central limitation of the current study. The IGT is only one measure, and although this measure is sensitive to many different processes involved in value-based decision making (as well as to some general aspects of intellectual functioning; see earlier discussion), it is not specific for any single decision process (Buelow & Sur, 2009). The decision literature distinguishes between

many fine-grained decision processes, including aversion to risk, aversion to ambiguity, aversion to loss, and the ability to learn from rewards and punishments (Pessiglione et al., 2006; Schönberg et al., 2007; Vaidya et al., 2013) . Changes in any of these processes could have effects on IGT performance. Indeed, different groups have attributed poor IGT performance to a deficit in reversal learning (Fellows, 2007; Fellows & Farah, 2005), a preference for taking risks (Dunn, Dalgleish, & Lawrence, 2006), or insensitivity to either rewards or punishments (Franken & Muris, 2005). Although it is possible that the impairments observed across different mental illnesses in the current study are all traceable back to the same underlying process within value-based decision making, it seems more likely that different mental illnesses affect different decision processes, but all of these effects lead to poorer performance on the IGT. Our findings therefore show that across mental illnesses, there is impairment within the broad class of processes involved in value-based decision making to which the IGT is sensitive, but these findings do not yet identify the specific processes within that broad class that are affected by specific disorders.

A very promising avenue of future research, then, would be to assess decision making in different mental illnesses by using a wider range of tasks that more clearly isolate specific decision processes. Studies could use a battery of tasks developed in the decision literature to assess the specific processes of risk aversion, ambiguity aversion, loss aversion, reward learning, and punishment learning. To focus on two of these constructs, we note that people might perform poorly on the IGT because they have a lower degree of risk aversion (Holt & Laury, 2002; Levy et al., 2010); that is, even once they know the probabilities and the outcomes associated with each deck, they are more willing to choose the higher-risk (i.e., higher-variance) disadvantageous decks. Alternatively, people might perform poorly on the IGT because they are slower to learn from rewards and punishments (Pessiglione et al., 2006; Schönberg et al., 2007; Vaidya et al., 2013); that is, it takes them a longer time to learn the probabilities and the outcomes associated with each deck. The decision tasks necessary to dissociate these two possibilities already exist.

In addition, performance on many of these more fine-grained decision tasks has been associated with specific neural systems. For example, neuroimaging studies using standard tasks to assess people's risk preferences have identified neural responses that scale with risk in the cingulate cortex, anterior insula, and inferior prefrontal cortex, and these neural responses predict an individual's degree of risk aversion (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009; Rudorf, Preuschoff, & Weber, 2012). In contrast, individual differences in reward learning have also been identified, and these are associated with neural signals in the ventral striatum that scale with reward prediction errors (Schönberg et al., 2007).

Therefore, a carefully selected battery of tasks, unlike the IGT alone, would be capable of identifying the specific "signature" of decision processes affected by a given mental illness. Comparison of these signatures across mental illnesses would then permit identification of their commonalities and distinguishing features and would furthermore lead to testable hypotheses about the neural systems affected by mental illness. Such investigations might start with those mental illnesses for which there are strong effects in the current meta-analyses, such as mood or personality disorders. This kind of study would fall squarely within the current push in clinical psychopathology toward transdiagnostic investigations of specific psychological processes. Although such a research effort would warrant a sizable investment, the current meta-analyses suggest that such an investment would be highly likely to yield interesting, informative results.

In this light, the current meta-analyses provide a broad "screen" for possible impairments in value-based decision making by assessing IGT performance across mental illnesses. The obvious next step, given the widespread impairments on the IGT that we documented, is to follow up on this screen to identify the specific decision processes that are impaired in specific disorders. Impaired decision making is generally not a focus in psychopathology. Indeed, the *DSM–5* considers decision-making impairment a possible symptom for only one disorder, major depressive disorder. The current meta-analyses suggest that further investigations of value-based decision making in mental illness, along the lines followed in the nascent field of computational psychiatry (Montague et al., 2012), hold substantial promise for the identification of specific decision processes that are adversely affected across disorders.

Author Contributions

D. Mukherjee designed the study, reviewed the literature, and collected and analyzed the data. D. Mukherjee and J. W. Kable wrote the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

References marked with an asterisk indicate studies included in the meta-analyses. Additional supporting information may be found at http://cpx.sagepub.com/content/by/supplemental-data

References

Ahn, W. Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical

Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, *4*, 95–110.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, *19*, 5473–5481.

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, *39*, 376–389.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*, 2189–2202.

Bickel, W. K., & Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. *Addiction*, *96*, 73–86.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive meta-analysis (Version 2) [Computer software]. Englewood, NJ: Biostat.

Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J., & Markowitsch, H. J. (2005). Decision-making deficits of Korsakoff patients in a new gambling task with explicit rules: Associations with executive functions. *Neuropsychology*, *19*, 267–277.

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology Review*, *19*, 102–114. doi:10.1007/s11065-009-9083-4

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences, USA*, *106*, 7745–7750.

Cavedini, P., Gorini, A., & Bellini, L. (2006). Understanding obsessive-compulsive disorder: Focus on decision-making. *Neuropsychology Review*, *16*, 3–15.

Cavedini, P., Riboldi, G., D'Annucci, A., Belotti, P., Cisima, M., & Bellodi, L. (2002). Decision-making heterogeneity in obsessive-compulsive disorder: Ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia*, *40*, 205–211.

Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive-compulsive disorder and trichotillomania. *Neuropsychologia*, *45*, 654–662.

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *Journal of Neuroscience*, *29*, 12574–12583.

Clark, L., Chamberlain, S. R., & Sahakian, B. J. (2009). Neurocognitive mechanisms in depression: Implications for treatment. *Annual Review of Neuroscience*, *32*, 57–74. doi:10.1146/annurev.neuro.31.060407.125618

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.

Davis, C. (2011). Unpublished data.

Dolan, S. L., Bechara, A., & Nathan, P. E. (2007). *Executive dysfunction as a risk marker for substance abuse: The role of impulsive personality traits*. Poster session (Paper 24) presented at the Samuel B. Guze Symposium on Alcoholism. Available at http://digitalcommons.wustl.edu/guzeposter2007/24

Dom, G., De Wilde, B. D., Hulstijn, W., & Sabbe, B. (2007). Dimensions of impulsive behavior in abstinent alcoholics. *Personality and Individual Differences*, *42*, 465–476.

Dom, G., De Wilde, B., Hulstijn, W., Van Den Brink, W., & Sabbe, B. (2006). Decision-making deficits in alcohol-dependent patients with and without comorbid personality disorder. *Alcoholism: Clinical and Experimental Research*, *30*, 1670–1677.

Dom, G., Sabbe, B., Hulstijn, W., & Van Den Brink, W. (2005). Substance use disorders and the orbitofrontal cortex: Systematic review of behavioural decision-making and neuroimaging studies. *British Journal of Psychiatry*, *187*, 209–220.

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews*, *30*, 239–271.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Fellows, L. K. (2007). The role of orbitofrontal cortex in decision-making. *Annals of the New York Academy of Sciences*, *1121*, 421–430.

Fellows, L. K., & Farah, M. J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex*, *15*, 58–63.

Follette, W. C. (1996). Introduction to the special section on the development of theoretically coherent alternatives to the *DSM* system. *Journal of Consulting and Clinical Psychology*, *64*, 1117–1119. doi:10.1037/0022-006X.64.6.1117

Franken, I. H. A., & Muris, P. (2005). Individual differences in decision-making. *Personality and Individual Differences*, *39*, 991–998.

Fridberg, D. J., Queller, S., Ahn, W. Y., Kim, W., Bishara, A. J., Busemeyer, J. R., & Stout, J. C. (2010). Cognitive mechanisms underlying risky decision-making in chronic cannabis users. *Journal of Mathematical Psychology*, *54*, 28–38.

Grecius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., . . . Schatzberg, A. F. (2007). Resting-state

functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, *62*, 429–437.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557–560.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2002). Somatic markers, working memory, and decision-making. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 341–353.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review*, *92*, 1644–1655.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*, 748–751.

Jameson, T. L., Hinson, J. M., & Whitney, P. (2004). Components of working memory and somatic markers in decision-making. *Psychonomic Bulletin & Review*, *11*, 515–520.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*, 1625–1633.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron*, *63*, 733–745.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Lawrence, N. S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., & Phillips, M. L. (2006). Decision-making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology*, *20*, 409–419.

Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, *103*, 1036–1047.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Maurex, L., Zaboli, G., Wiens, S., Åsberg, M., Leopardi, R., & Öhman, A. (2009). Emotionally controlled decision-making and a gene variant related to serotonin synthesis in women with borderline personality disorder. *Scandinavian Journal of Psychology*, *50*, 5–10.

Mayberg, H. S. (2006). Defining neurocircuits in depression: Insight from functional neuroimaging studies of diverse treatments. *Psychiatric Annals*, *36*, 259–268.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*, 72–80.

Murphy, F. C., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., Paykel, E. S., & Sahakian, B. J. (2001). Decision-making cognition in mania and depression. *Psychological Medicine*, *31*, 679–693.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042–1045.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, *57*, 119–127.

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, *20*, 262–270.

Rilling, J. K., King-Casas, B., & Sanfey, A. G. (2008). The neurobiology of social decision-making. *Current Opinion in Neurobiology*, *18*, 159–165.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Rudorf, S., Preuschoff, K., & Weber, B. (2012). Neural correlates of anticipation risk reflect risk preferences. *Journal of Neuroscience*, *32*, 16683–16692.

Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinssen, R. K., . . . Cuthbert, B. N. (2010). Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology*, *119*, 631–639. doi:10.1037/a0020909

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*, 12860–12867.

Sevy, S., Burdick, K. E., Visweswaraiah, H., Abdelmessih, S., Lukin, M., Yechiam, E., & Bechara, A. (2007). Iowa Gambling Task in schizophrenia: A review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophrenia Research*, *92*, 74–84.

Tolin, D. F., Abramowitz, J. S., Brigidi, B., & Foa, E. B. (2003). Intolerance of uncertainty in obsessive-compulsive disorder. *Journal of Anxiety Disorders*, *17*, 233–242.

Vaidya, J. G., Knutson, B., O'Leary, D. S., Block, R. I., & Magnotta, V. (2013). Neural sensitivity to absolute and relative anticipated reward in adolescents. *PLoS ONE*, *8*(3): e58708. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0058708

Verdejo-García, A., & Bechara, A. (2009). A somatic marker theory of addiction. *Neuropharmacology*, *56*, 48–62.

Verdejo-García, A., Pérez-García, M., & Bechara, A. (2006). Emotion, decision-making and substance dependence: A somatic-marker model of addiction. *Current Neuropharmacology*, *4*, 17–31.

Volkow, N. D., & Fowler, J. S. (2000). Addiction, a disease of compulsion and drive: Involvement of the orbitofrontal cortex. *Cerebral Cortex*, *10*, 318–325.

Weller, J. A., Levin, I. P., & Bechara, A. (2010). Do individual differences in Iowa Gambling Task performance predict adaptive decision-making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology*, *32*, 141–150.