Title: Deciphering reward-based decision-making in schizophrenia: a meta-analysis and behavioral modeling of the Iowa Gambling Task

Author names: Linda T. Betz¹, Paolo Brambilla², Andrej Ilankovic³, Preethi Premkumar⁴, Myung-Sun Kim⁵, Stéphane Raffard⁶⁷, Sophie Bayard⁷, Hikaru Hori⁸, Kyoung-Uk Lee⁹, Seung Jae Lee¹⁰, Nikolaos Koutsouleris¹, Joseph Kambeitz¹

Author affiliations: ¹Department of Psychiatry, Ludwig-Maximilian-University Munich, Munich, Germany, linda.betz@gmx.de, nikolaos.koutsouleris@med.uni-muenchen.de, joseph.kambeitz@med.uni-muenchen.de; ²Scientific Institute IRCCS “E. Medea”, Bosisio Parini, Lecco, Italy, paolo.brambilla1@unimi.it; ³Psychiatry Clinic, Clinical Center of Serbia, University of Belgrade, Belgrade, Serbia, andrejilankovic@gmail.com; ⁴Department of Psychology, School of Social Sciences, Nottingham Trent University, Nottingham, UK, preethi.premkumar@ntu.ac.uk; ⁵Department of Psychology, Sungshin Women’s University, Seoul, Republic of Korea, kimms@sungshin.ac.kr; ⁶University Department of Adult Psychiatry, La Colombière Hospital, CHRU Montpellier, Montpellier, France, s-raffard@chu-montpellier.fr; ⁷Laboratoire Epsylon, EA 4556, Université Paul Valéry Montpellier 3, Montpellier, France, sophie.bayard@univ-montp3.fr; ⁸Department of Psychiatry, University of Occupational and Environmental Health, 1-1 Iseigaoka, Yahatanishi-ku, Kitakyushu, Fukuoka 8078555, Japan, hori-h@med.uoeh-u.ac.jp; ⁹Department of Psychiatry, Uijeongbu St. Mary's Hospital, The Catholic University of Korea, College of Medicine, Uijeongbu, Korea, mindcure@catholic.ac.kr; ¹⁰Department of Psychiatry, School of Medicine, Kyungpook National University, Daegu, Korea, sjl7670@hotmail.com

Address of corresponding author: Linda Betz, Department of Psychiatry and Psychotherapy, Nußbaumstraße 7, 80336 Munich, Germany; Phone: +49 89 4400 55880; Fax: +49 89 4400 54749; e-mail: linda.betz@med.uni-muenchen.de
Abstract

Background: Patients with schizophrenia (SZP) have been reported to exhibit impairments in reward-based decision-making, but results are heterogeneous with multiple potential confounds such as age, intelligence level, clinical symptoms or medication, making it difficult to evaluate the robustness of these impairments.

Methods: We conducted a meta-analysis of studies comparing the performance of SZP and healthy controls (HC) in the Iowa Gambling Task (IGT) as well as comprehensive analyses based on subject-level data (n = 303 SZP, n = 188 HC) to investigate reward-based decision-making in SZP. To quantify differences in the influence of individual deck features (immediate gain, gain frequency, net loss) between SZP and HC, we additionally employed a least-squares model.

Results: SZP showed statistically significant suboptimal decisions as indicated by disadvantageous deck choices (d from 0.51 to -0.62) and lower net scores (d from -0.35 to -1.03) in a meta-analysis of k = 29 samples (n = 1127 SZP, n = 1149 HC) and these results were confirmed in a complementary subject-level analysis. Moreover, decision-making in SZP was characterized by a relative overweighing of immediate gain and net losses and an underweighting of gain frequency. Moderator analyses revealed that in part, decision-making in the IGT was moderated by intelligence level, medication and general symptom scores.

Conclusion: Our results indicate robust impairments in reward-based decision-making in SZP and suggest that decreased cognitive resources, such as working memory, may contribute to these alterations.

Keywords: schizophrenia, decision-making, reward, Iowa Gambling Task, meta-analysis, linear modeling
1. Introduction

Patients with schizophrenia (SZP) exhibit deficits across a wide range of measures of executive control/working memory (WM; Collins et al., 2014) and reward-processing/reinforcement learning (RL; Gold et al., 2008; Juckel et al., 2006; Strauss et al., 2012; Waltz et al., 2007). A particularly debilitating aspect of the illness potentially associated with these impairments is maladaptive decision-making, contributing significantly to several functional deficits, such as poor treatment adherence, financial difficulties and interpersonal conflicts (Fond et al., 2013).

One popular experimental paradigm to examine decision-making under ambiguity in SZP is the Iowa Gambling Task (IGT; Bechara et al., 1994). On each trial, participants choose a card from one of four decks and receive a monetary gain or loss. The decks differ in three properties: the amount of immediate gain, the relative frequency of gains vs. losses (gain frequency) and the relative number of net losses, i.e., the number of instances when the sum of gains and losses on a card is below zero (net losses). The combination of these features is unique for each deck and unknown to the participants (supplementary tables 1-2). The goal is to maximize monetary outcome through adaptive decision-making in 100 trials: two decks (A, B) are disadvantageous; two decks (C, D) are advantageous.

To disentangle the factors underlying IGT performance, computational cognitive models such as the Expectancy Valence Learning (EVL) model have been employed (Busemeyer and Stout, 2002) to quantify latent cognitive components (attention to gains, learning rate, recency, response consistency) and a least-squares model has been used to identify the relative importance of the deck features (supplementary tables 1-2) in decision-making (Horstmann et al., 2012).

Previous research has evidenced the important role of high-level cognitive functions, especially memory (WM), for successful performance in the IGT (Bagneux et al., 2013;
Demaree et al., 2010; Hawthorne and Pierce, 2015; Maia and McClelland, 2004; Stocco et al., 2009). In SZP, impaired WM is arguably a core symptom related to prefrontal abnormalities, as supported by behavioral observations (Anticevic et al., 2011; Collins et al., 2014; Park et al., 1999; Strauss et al., 2012), consistent hypofrontality during WM tasks (Glahn et al., 2005) and alterations of prefrontal D1 receptor transmission involved in WM deficits (Abi-Dargham et al., 2002).

Importantly, WM interacts with basic processes of reinforcement learning (RL) in learning paradigms (Collins et al., 2017), making the dopaminergic system – implicated in reward processing (Pizzagalli et al., 2008; Schultz et al., 1997; St Onge and Floresco, 2009) – another factor in IGT performance. Deficient task performance may therefore also be linked to well-documented alterations in the dopaminergic system of SZP (Carlsson, 1988), specifically increased striatal dopamine synthesis capacity (Howes et al., 2012) interfering with reward prediction errors (Juckel et al., 2006).

A previous meta-analysis of eight studies revealed that SZP showed poor performance compared to HC in the IGT, potentially stemming from a RL deficit (Brown et al., 2015). However, only deck choices were meta-analyzed and moderator variables were not accounted for. Lastly, the specific pattern of suboptimal deck choices in schizophrenia was not examined formally, e.g. with a model. Therefore, we employed a meta-analytic approach to integrate a larger body of available evidence to evaluate the differences in decision-making between SZP and HC in several outcome measures of the IGT: block net scores, deck choices, and EVL model parameters. Our secondary goal was to investigate the moderating effects of demographical and clinical variables on the difference in IGT performance between HC and SZP. Third, to validate the results from the meta-analysis, we used original, subject-level data for modeling decision-making behavior of SZP and HC in the IGT. Lastly, we modeled the influence of the statistical properties of the decks (frequency and magnitude of
gains/losses; net losses) on decision-making in SZP relative to HC to gain insight into the nature of their deficits on the IGT (Horstmann et al., 2012).

2. Methods

2.1 Literature Search and Data Extraction

Following recommended guidelines as defined in the PRISMA statement (Moher et al., 2009), we conducted a systematic literature search in the databases PubMed, PsycINFO, and Web of Science in November 2017 using the search term ("psychosis" OR "psychotic" OR "schizophrenia" OR "schizophrenic") AND ("Iowa Gambling task" OR "Gambling task" OR "Bechara Gambling task"), searching for studies published between 1990/01/01 and 2017/11/15. Reference lists of published reviews and studies were used to identify additional publications. To be included, a study had to (1) include a sample of SZP or schizoaffective disorder as defined by the DSM IV, (2) include a sample of healthy control participants matched for age (case-matched or no significant differences in mean age between the groups), (3) report results from the IGT in a way that sufficient information could be extracted to calculate at least one of the effect sizes of interest (net score in each of the five blocks, total number of cards selected from each of the individual decks, parameters from computational modeling). From all included studies, we extracted means, standard deviations and group sizes to calculate the standardized mean difference (SMD, expressed as $d$; Hedges, 1981) between SZP and HC for block net scores, deck choices and/or model parameters, as well as moderator variables (year of publication, continent where the study was conducted, quality of the study (derived from an adapted version of the Newcastle-Ottawa Scale, Wells et al., 2000, supplementary table 3), demographic information of the control and patient sample (age, gender, years of education, intelligence quotient (IQ) levels), and clinical characteristics of the patient sample (illness duration, age of onset, symptom ratings [Positive and Negative Symptom Scale (PANSS); Kay et al., 1987], amount of current neuroleptic
medication [chlorpromazine equivalents (CPZE)], percent of SZP medicated with antipsychotics, percent of SZP medicated with first-generation antipsychotics) (supplementary methods for details on data extraction).

2.3 Statistical Analysis

2.3.1 Meta-analysis. Psychiatric populations are typically heterogeneous due to differences in illness duration, previous medication and symptoms. To account for this heterogeneity, we employed random-effects meta-analysis models. We inspected funnel plots and tested for funnel plot asymmetry using Egger’s test (Egger et al., 1997). If Egger’s test was significant, we used the trim-and-fill method to estimate the number of missing studies and correct the estimated effect size (Duval and Tweedie, 2000a, 2000b). When important heterogeneity was identified, we used random effects univariate meta-regression models to examine whether any of the extracted variables described above were associated with performance differences between SZP and HC.

2.3.2 Behavioral modeling. To increase statistical power compared to the meta-analysis based on summary data, we fitted linear mixed-effects models to a subset of original, patient-level data to investigate differences in decision-making between HC and SZP for each of the IGT-outcomes of interest (block net scores, deck choices), pooling available data and including study as a random effect. We also investigated moderating effects of the demographic and clinical variables detailed above (supplementary methods 1).

2.3.3 Least-Squares modeling. We modeled decision-making on the IGT with a system of linear equations, as \( \mathbf{Xb} = \mathbf{y} \) (Horstmann et al., 2012). Least-squares provides a solution \( \mathbf{b} \) for the system, describing a linear relationship between the statistical properties of the decks (supplementary tables 1-2) included in \( \mathbf{X} \) and the deck choices in \( \mathbf{y} \) (supplementary methods 2). We applied this model to aggregated and subject-level deck choices. To explore moderation effects on the model parameters, we used linear mixed-effects models.
We conducted all analyses in the R statistical language for computing, version 3.4.2 (R Core Team, 2017) using the packages ‘metafor’ (Viechtbauer, 2010), ‘lme4’ (Bates et al., 2015) and ‘lmerTest’ (Kuznetsova et al., 2015). We obtained p-values for the coefficients in the linear mixed-effects models based on the Satterthwaite approximation for denominator degrees of freedom implemented in lmerTest. A significance level of \( p = .05 \) was considered throughout.

2. Results

Twenty-nine samples described in 27 publications [two articles (Raffard et al., 2011; Sevy et al., 2007) reported data from two samples each] met our inclusion criteria and were included in the analyses (figure 1; included studies in supplementary table 4; study/sample characteristics and moderator variables in supplementary table 5). In total, this comprised \( n = 1127 \) SZP and \( n = 1149 \) HC.

3.1 Meta-analyses and meta-regression

Table 1 provides a summary of the meta-analysis results for all IGT indices.

3.1.1 Block net scores. Twenty-five samples from 23 publications (two studies (Raffard et al., 2011; Sevy et al., 2007) described data from two samples each) with a total of \( n=965 \) SZP and \( n=921 \) HC reported block net scores. Meta-analyses indicated significant reductions of net scores in SZP in block 2 \((d=-0.34, 95\%-CI: -0.51 \text{ to } -0.18, p < .001, I^2 = 66.2\%)\), block 3 \((d=-0.70, -0.96 \text{ to } -0.44, p < .001, I^2=85.1\%)\), block 4 \((d=-0.94, -1.25 \text{ to } -0.63, p < .001, I^2=89.5\%)\) and block 5 \((d=-1.06, -1.50 \text{ to } -0.63, p < .001, I^2=94.5\%); figure 3) as compared to HC, but not in block 1 \((d=0.09, -0.04 \text{ to } 0.23, p=.154, I^2=44.7\%); figure 2A\) and supplementary figures 1A-5). Egger’s test was significant for block 5 \((z=-2.81, p=.005)\) and trim-and-fill analysis suggested that six studies were potentially ‘missing’ on the right-hand side of the funnel plot \((p=.008, \text{ supplementary figure 6})\). After imputing these putatively missing studies, the mean effect size was -0.58 \((-1.13 \text{ to } -0.04, p=.035, I^2=97.0\%); Meta-
regression indicated that later mean age of illness onset was associated with a smaller
difference between SZP and HC in block 5 net scores (number of studies that reported the
moderator variable \( k=10, \beta=0.15 \), Standard Error (SE)=0.06, \( p=.016 \)). This finding was
mainly driven by the early-onset schizophrenia sample of Kester et al. (2006). When
excluded, the moderator effect was no longer significant (supplementary results).

3.1.2 Deck choices. Seventeen studies from 16 publications (one study (Sevy et al.,
2007) described data from two samples) comprising \( n=648 \) SZP and \( n=566 \) HC reported deck
choices. Meta-analyses indicated significant increases in the number of cards chosen from
deck A (\( d=0.35, 0.21 \) to 0.49, \( p<.001, F=25.54\% \)) and deck B (\( d=0.51, 0.29 \) to 0.71, \( p<
.001, F=68.19\% \), figure 4) in SZP compared to HC. Conversely, SZP drew significantly less
cards from deck D (\( d=-0.62, -0.84 \) to -0.41, \( p<.001, F=66.13\% \)) than HC. Choices from
deck C were not significantly different between the groups (\( d=-0.13, -0.37 \) to 0.11, \( p=.278,
F=73.92\% \)) (figure 2B and supplementary figures 7-9).

Meta-regression indicated that higher level of education in the control sample was
associated with increased choices from deck A in SZP compared to HC (\( k=16,
\beta=0.08, SE=0.04, p=.029 \)). Additionally, there was evidence for a relative decrease in choices
from deck A in SZP compared to HC with increasing IQ (\( k=11, \beta=-0.04, SE=0.02, p=.009 \)),
higher positive (\( k=9, \beta=-0.05, SE=0.02, p=.04 \)), higher general psychopathology (\( k=7, \beta=-
0.02, SE=0.01, p=.04 \)) and higher total PANSS scores in the patient sample (\( k=11, \beta=-
0.009, SE=0.004, p=.024 \)). Similarly, lower rates of antipsychotic-medicated patients in the
sample were associated with a decrease in choices from deck A (\( k=5,
\beta=0.58, SE=0.20, p=.004 \)) and an increase in cards chosen from deck B (\( k=5, \beta=-
0.92, SE=0.37, p=.012 \)). These effects were driven by the Zhang et al. (2015) study. When
excluded, the moderator effects were no longer significant (supplementary results).
3.1.3 EVL model parameters. Five samples from four publications (one study (Sevy et al., 2007) included two samples) reported EVL model parameters \((n=187\) SZP and \(n=230\) HC). Meta-analyses indicated no significant differences between SZP and HC in any of the model parameters (attention to gains/losses \((d=0.06, -0.15\) to \(0.26, p=.597, I^2=0.0\%\)), response consistency \((d=-0.01, -0.47\) to \(0.44, p=.958, I^2=73.75\%\)), learning rate \((d=0.15, -0.32\) to \(0.62, p=.524, I^2=75.13\%\)) (figure 2C and supplementary figures 10-12). Egger’s test was significant for response consistency \((z=4.05, p<.001)\) and learning rate \((z=-3.97, p<.001)\). Trim-and-fill analyses detected no ‘missing’ studies in both cases \((p=.50; supplementary figures 13-14)\). Due to the small number of studies, we do not report moderator analyses in this section.

3.2 Modeling based on individual subject data

Original individual subject data comprised \(n=491\) subjects (303 SZP and 188 HC) provided by authors of five studies included in the meta-analysis (Hori et al., 2014; Kim et al., 2009; Kim and Kang, 2016; Premkumar et al., 2008; Raffard et al., 2011).

3.2.1 Behavioral modeling of block net scores. Block net scores were available for all \(n=491\) subjects. In line with our meta-analysis, linear mixed-effect models indicated that diagnosis was associated with lower block net scores in SZP compared to HC during block 2 \((\beta=-1.90, SE=0.75, p=.026)\), block 3 \((\beta=-4.36, SE=0.82, p<.001)\), block 4 \((\beta=-6.00, SE=0.96, p<.001)\), and block 5 \((\beta=-7.00, SE=1.13, p<.001)\), but not during block 1 \((p = .722)\) (supplementary figure 15A). In \(n=391\) subjects (228 SZP, 163 HC), we examined the role of moderator effects (gender, age, IQ) on block net scores. Analogously, we examined these moderator effects in \(n=207\) subjects (125 SZP, 82 HC) for whom deck choices were available. Results indicated an interaction effect of diagnosis and IQ scores on block 2 net scores in patients and controls: IQ scores were positively correlated with block net scores in HC, while this relationship was attenuated in patients \((\beta=-0.14, SE=0.07, p=.038)\). Higher net
scores in block 3 were associated with higher IQ scores in both groups ($\beta=0.11$, SE=0.06, $p=.049$). In data from $n=228$ SZP, higher doses of medication were associated with significantly decreased net scores during block 1 ($\beta=-0.76$, SE 0.37, $p = .042$), block 2 ($\beta=-0.90$, SE=0.43, $p=.041$) and at non-significant trend-level during block 3 ($\beta=-0.78$, SE=0.45, $p=.085$). Additionally, there was a trend suggesting that higher PANSS general symptom scores were associated lower net scores in block 4 ($\beta=-0.15$, SE=0.08, $p=.073$).

3.2.2 Behavioral modeling of deck choices. Deck choices were available for $n=207$ subjects (125 SZP, 82 HC). In accordance with the meta-analyses, linear mixed-effect models indicated that SZP selected significantly more cards from deck A ($\beta=-2.22$, SE=1.00, $p = .027$) and B ($\beta=5.93$, SE=1.49, $p < .001$) than HC, but less cards from deck C ($\beta=-5.95$, SE=1.24, $p < .001$). Results for deck D were not significant ($p=.126$). (supplementary figure 15B). Moderator analyses showed that IQ scores differentially affected choices from deck B: higher IQ scores resulted in decreased choices from deck B in HC, but this effect was attenuated in SZP ($\beta=0.39$, SE=0.14, $p=.009$). A linear mixed-effect model based on data from $n=125$ SZP identified a non-significant trend indicating that higher PANSS general symptom scores were linked to fewer cards selected from deck C ($\beta=-0.16$, SE = 0.09, $p=.069$).

3.3 Least-Squares modeling

Linear mixed-effects models based on aggregated data indicated an increase in immediate gain in patients relative to controls ($\beta=5.87$, SE=0.81, $p < .001$). Patients gave less weight to gain frequency ($\beta=-1.93$, SE=0.83, $p=.033$), and net loss ($\beta=-2.79$, SE=0.67, $p < .001$). Results from subject-level data were similar for immediate gains ($\beta=-4.48$, SE=1.36, $p=.001$), but different for gain frequency ($\beta=2.34$, SE=0.90, $p=.010$) and net loss ($\beta=0.13$, SE=1.15, $p = .907$) (figure 5).
The moderator analyses conducted on subject-level data revealed a significant interaction of IQ scores and diagnosis that moderated immediate gain ($\beta=0.49$, SE=0.14, $p < .001$), gain frequency ($\beta=-0.24$, SE=0.09, $p=.011$), indicating that in HC, higher IQ scores resulted in less weighting of immediate gain and increased weighting of gain frequency, while these relationships were attenuated or reversed in patients. Additionally, we observed non-significant trends indicating that the weighting of gain frequency decreased with higher negative symptoms ($\beta=-0.17$, SE=0.10, $p=.093$) and that the weighting of immediate gains was increased with higher general symptom scores ($\beta=0.22$, SE=0.13, $p=.096$) in SZP.

3. Discussion

We present a meta-analysis of $k=29$ samples from 27 publications ($n=1127$ SZP, $n=1149$ HC) as well as a comprehensive analysis of subject-level data ($n=303$ SZP and $n=188$ HC). Our results demonstrate impaired decision-making in SZP compared to HC in the IGT with small to large effect sizes, depending on the IGT index. HC maximized net scores, displaying a steep learning curve, while SZP did not improve during the task. Across studies, patients chose significantly more cards from the momentary profitable, but overall disadvantageous decks, whereas HC selected significantly more cards from the advantageous deck D. Importantly, these results were robust with respect to potential publication biases and the inclusion of confounding factors (including year of publication, study quality, continent where the study was conducted) and the results could be confirmed with subject-level data.

Findings from the meta-analysis revealed a pattern of deck selections that is specific to schizophrenia: while SZP and controls selected deck C with similar frequency, SZP selected decks A and B more frequently and deck D less frequently than HC. Deck B is associated with frequent high gains and rare, but very large losses, resulting in an overall disadvantageous outcome. Despite this negative long-term balance, HC also frequently pick deck B (Dunn et al., 2006; Lin et al., 2007). In line with previous literature (Horstmann et al.,
2012), the least-squares model showed that decision-making in HC is primarily driven by
gain frequency as indicated by a relative preference for decks with high-frequency gains (B
and D). Additionally, HC distinguish between the decks with low-frequency gains by
weighting net losses negatively: they choose deck C (which never yields a net loss) over deck
A (which yields frequent net losses). In contrast, SZP decision-making was mainly driven by
net losses and immediate gains. Most interestingly, patients seem to prefer deck B by a large
margin over all other decks for other reasons than controls: while controls are attracted by the
high gain frequency associated with this deck, patients are drawn to deck B due to a
combined influence of low net losses and high immediate gains. These findings might
illustrate how patients disregard long-term outcome in their decision-making and focus on
high immediate gains. This pattern of decision-making is in keeping with reported intact
sensitivity to immediate and reliable rewards in schizophrenia, as opposed to more complex
or temporally remote rewards (Heerey et al., 2008; Juckel et al., 2006; Waltz et al., 2007).

Past research has indicated that the IGT is a cognitively demanding task requiring high-
level cognitive functions such as holding the experimental contingencies in WM (Bagneux
et al., 2013; Demaree et al., 2010; Hawthorne and Pierce, 2015; Maia and McClelland, 2004;
Stocco et al., 2009). Integration of information across decks and trials is particularly
important to capture gain frequency since this parameter differs between decks A/C and
decks B/D. Net losses, by contrast, can be incorporated into decision-making relatively easily
since deck A is the only deck with markedly high net losses, obviously differing from all
other decks. Similarly, capturing high immediate gains does not require active maintenance
of information. Thus, the pattern of decision-making found in SZP may suggest that they rely
more on deck features that pose less load on their WM system. These parameters may be
captured by integrating outcome magnitudes using trial-by-trial RL processes, which seem to
be relatively spared in schizophrenia (Collins et al., 2017; Heerey et al., 2008). Importantly,
even patients with comparatively intact cognitive resources do not show altered decision-making strategies: contrary to HC, patients with higher IQ scores do not change their weighing in favor of gain frequency and against immediate gains. It has been speculated that such a pattern of behavior mirrors patients’ reduced confidence in their WM capacities (Collins et al., 2014). Alternatively, PFC-dependent motivational factors, such as willingness to expend effort, may influence patients’ decision-making (Barch et al., 2014). Given that no direct measure of WM nor motivation was assessed in this meta-analysis, the determinants of the specific pattern of IGT-performance warrant further investigation.

Our data suggest that decision-making in the IGT is negatively influenced by high doses of neuroleptic medication, especially during the initial blocks when intact WM functioning seems to be critically involved in learning the deck contingencies (Hawthorne and Pierce, 2015; Horstmann et al., 2012; Stocco et al., 2009). Chronic dopamine blockage may exacerbate WM dysfunction in schizophrenia as proposed by the inverted U-shaped relationship between dopamine levels and WM function (Cools and D'Esposito, 2011). The use of the deck feature gain frequency, which is strongly dependent on WM and central to successful task performance in controls, may be particularly affected by antipsychotics. Conversely, trial-by-trial aspects of RL associated with striatal-dopaminergic signaling that seem to be normalized by antipsychotic medication (Insel et al., 2014) may only be secondary to succeeding in the IGT. Such effects, however, are difficult to distinguish from general effects of illness severity on task performance also suggested by the data.

4. Limitations

The limited range of some of the moderator variables may have contributed to their relative nonsignificance in the present analyses. The patient samples analyzed were relatively homogeneous in terms of stability and chronicity of the illness, age of onset and type and amount of antipsychotic medication, and data on treatment duration was largely unavailable
across studies. This makes it difficult to assess in more detail the effects of disease progression or duration and type of antipsychotic medication known to influence the performance on neuropsychological tests associated with IGT performance (Premkumar et al., 2008; Yip et al., 2009).

Additionally, some of the observed moderator effects were driven by single outlier studies. The early-onset schizophrenia sample (Kester et al., 2006) showed particularly poor task performance, potentially driven by more pronounced WM dysfunction in this population compared to adult-onset schizophrenia (Frangou, 2009). Similarly, the only available sample of drug-naïve patients (Zhang et al., 2015) showed altered decision-making by means of a strong focus on decks with high immediate gains. However, drug-naïvety co-occurred with high symptomatology and intelligence in this sample, complicating the evaluation of the role of medication. Thus, conclusions on these moderator effects should be tentative and substantiated by further research.

The aggregation of data inherent to the process of meta-analysis leads to a loss of information and potentially imprecise results. We tried to reduce this risk by validating our results with the help of original data. Results could be confirmed overall, but discrepancies in deck choices and least-squares parameters emerged. Subject-level analyses of these outcomes were based on data from two studies only, compromising the generalizability of these results. Especially, behavior of HC reported in these studies differed from the aggregated findings in the meta-analysis. Varying eligibility criteria for HC across studies may account for some of the heterogeneity in IGT performance observed in our and previous analyses (Steingroever et al., 2013).

Moreover, even in the analyses at subject-level, data was aggregated per block and deck. This may have led to an underestimation of the importance of aspects of RL that are primarily reflected in changes in deck choices from trial to trial of participants, e.g. prolonged
sampling across or adaptive switching between decks. Using trial-by-trial data, more detailed cognitive modeling is possible: Recent computational cognitive models like the Prospect Valence Learning (PVL) model have revised details for utility, learning and choice probability functions and can model complex decision-making more accurately than the EVL model (Ahn et al., 2008; Fridberg et al., 2010). Similarly, modeling of changes of weighting of deck features across blocks may provide valuable insights into the adaptation of decision-making of SZP (Horstmann et al., 2012).

5. Conclusion and clinical implications

The deficits of SZP in the IGT were demonstrated in meta-analyses across 29 samples from 27 publications and overall confirmed by comprehensive modeling of individual subject data. A least-squares model illustrated that these deficits are primarily driven by an impaired integration of information about the relative frequency of gains vs. losses and an over-weighting of immediate gains suggesting that a WM deficit may contribute to suboptimal task performance in schizophrenia. This pattern of decision-making may relate to several behavioral patterns that pose problems in the treatment of SZP, such as poor medication adherence or substance abuse (Pristach and Smith, 1990). Therapeutic interventions should account for this decision-making behavior neglecting long-term outcome, e.g. by imparting strategies to increase the awareness of long-term consequences of actions (Evans et al., 2005). Additionally, assuring medication delivery using long-acting injectable medication can help to optimize treatment success in SZP (Barnes and Cursori, 1994).

Figure Legends:

Figure 1. Overview of the different stages of the systematic literature search according to the PRISMA guidelines (Moher et al., 2009).
Figure 2. Results from the meta-analysis. Weighted averages (by sample size) for patients with schizophrenia (SZP) and healthy controls (HC). (A) Net scores during the five blocks of the Iowa Gambling Task. (B) Mean number of cards chosen from each deck. (C) Parameters from the Expectancy-Valence (EVL) model. Atten = Attention to Gains versus Losses, Cons = Response Consistency, Learn = Learning Rate/Recency Parameter. Error bars represent the 95% confidence interval. ***p < .001.

Figure 3. Forest plot of the standardized mean difference between schizophrenic patients (SZP) and healthy controls (HC) in block 5 net score of the Iowa Gambling Task. Annotation: Raffard et al., 2011a: high insight SZP, Raffard et al., 2011b: low insight SZP; Sevy et al., 2007a: SZP without cannabis use disorder, Sevy et al., 2007b: SZP with cannabis use disorder.

Figure 4. Forest plot of the standardized mean difference between patients with schizophrenia (SZP) and healthy controls (HC) in the number of cards chosen from deck B in the Iowa Gambling Task. Annotation: Sevy et al., 2007a: SZP without cannabis use disorder, Sevy et al., 2007b: SZP with cannabis use disorder.

Figure 5. Results from least-squares modeling for patients with schizophrenia (SZP) and healthy controls (HC). (A) Weighted mean (by sample size) of the weights of least-squares modeling parameters. The model was applied to the aggregated deck choices reported in the studies included in the meta-analysis. (B) Mean weight of the parameters derived from applying the least-squares model to subject-level data. Error bars represent the 95% confidence interval. *p < .05, **p < .01
Table:

**Table 1.** Summary table for meta-analysis results for all IGT indices.

<table>
<thead>
<tr>
<th>IGT index</th>
<th>Effect Sizes</th>
<th>k Studies</th>
<th>n (SZP)</th>
<th>n (HC)</th>
<th>SMD</th>
<th>95%-CI</th>
<th>z</th>
<th>p</th>
<th>F (in %)</th>
<th>Egger's test</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Net Score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 1</td>
<td>25</td>
<td>23</td>
<td>965</td>
<td>921</td>
<td>0.09</td>
<td>[-0.04, 0.23]</td>
<td>1.43</td>
<td>.15</td>
<td>44.7</td>
<td>0.05</td>
<td>0.957</td>
<td></td>
</tr>
<tr>
<td>Block 2</td>
<td>25</td>
<td>23</td>
<td>965</td>
<td>921</td>
<td>-0.34</td>
<td>[-0.51, -0.18]</td>
<td>-4.01</td>
<td>&lt;.001</td>
<td>66.2</td>
<td>-0.60</td>
<td>.550</td>
<td></td>
</tr>
<tr>
<td>Block 3</td>
<td>25</td>
<td>23</td>
<td>965</td>
<td>921</td>
<td>-0.70</td>
<td>[-0.96, -0.44]</td>
<td>-5.30</td>
<td>&lt;.001</td>
<td>85.1</td>
<td>-0.30</td>
<td>.764</td>
<td></td>
</tr>
<tr>
<td>Block 4</td>
<td>25</td>
<td>23</td>
<td>965</td>
<td>921</td>
<td>-0.94</td>
<td>[-1.25, -0.63]</td>
<td>-5.89</td>
<td>&lt;.001</td>
<td>89.5</td>
<td>-1.03</td>
<td>.305</td>
<td></td>
</tr>
<tr>
<td>Block 5</td>
<td>25</td>
<td>23</td>
<td>965</td>
<td>921</td>
<td>-1.06</td>
<td>[-1.50, -0.63]</td>
<td>-4.77</td>
<td>&lt;.001</td>
<td>94.5</td>
<td>-2.81</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td><strong>Deck Choice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deck A</td>
<td>17</td>
<td>16</td>
<td>648</td>
<td>566</td>
<td>0.35</td>
<td>[0.21, 0.49]</td>
<td>4.91</td>
<td>&lt;.001</td>
<td>25.5</td>
<td>1.92</td>
<td>.054</td>
<td></td>
</tr>
<tr>
<td>Deck B</td>
<td>17</td>
<td>16</td>
<td>648</td>
<td>566</td>
<td>0.51</td>
<td>[0.29, 0.71]</td>
<td>4.53</td>
<td>&lt;.001</td>
<td>68.2</td>
<td>0.43</td>
<td>.667</td>
<td></td>
</tr>
<tr>
<td>Deck C</td>
<td>17</td>
<td>16</td>
<td>648</td>
<td>566</td>
<td>-0.13</td>
<td>[0.37, 0.11]</td>
<td>-1.08</td>
<td>.278</td>
<td>73.9</td>
<td>-0.15</td>
<td>.879</td>
<td></td>
</tr>
<tr>
<td>Deck D</td>
<td>17</td>
<td>16</td>
<td>648</td>
<td>566</td>
<td>-0.62</td>
<td>[-0.84, -0.41]</td>
<td>-5.69</td>
<td>&lt;.001</td>
<td>66.1</td>
<td>-0.96</td>
<td>.337</td>
<td></td>
</tr>
<tr>
<td><strong>EVL-Model Parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attention to gains/losses</td>
<td>5</td>
<td>4</td>
<td>187</td>
<td>230</td>
<td>0.06</td>
<td>[-0.15, 0.26]</td>
<td>0.53</td>
<td>.597</td>
<td>0.0</td>
<td>0.53</td>
<td>.599</td>
<td></td>
</tr>
<tr>
<td>response consistency</td>
<td>5</td>
<td>4</td>
<td>187</td>
<td>230</td>
<td>-0.01</td>
<td>[-0.47, 0.44]</td>
<td>-0.05</td>
<td>.958</td>
<td>73.8</td>
<td>4.05</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>learning rate</td>
<td>5</td>
<td>4</td>
<td>187</td>
<td>230</td>
<td>0.15</td>
<td>[-0.32, 0.62]</td>
<td>0.64</td>
<td>0.525</td>
<td>75.1</td>
<td>-3.97</td>
<td>&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

*Abbreviations:* IGT = Iowa Gambling Task, N = Number, SZP = Patient with schizophrenia, HC = Healthy control, SMD = Standardized Mean Difference, CI = Confidence Interval, z = z-value, p = p-value, EVL = Expectancy-Valence Learning.
<table>
<thead>
<tr>
<th>Author(s) and Year</th>
<th>Standardized Mean Difference [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ritter et al., 2004</td>
<td>0.01 [-0.66, 0.68]</td>
</tr>
<tr>
<td>Evans et al., 2005</td>
<td>-0.02 [-0.66, 0.61]</td>
</tr>
<tr>
<td>Rodríguez-Sánchez et al., 2005</td>
<td>-0.03 [-0.50, 0.45]</td>
</tr>
<tr>
<td>Shurman et al., 2005</td>
<td>-1.02 [-1.74, -0.29]</td>
</tr>
<tr>
<td>Kester et al., 2006</td>
<td>-3.31 [-4.27, -2.34]</td>
</tr>
<tr>
<td>Lee et al., 2007</td>
<td>-0.91 [-1.49, -0.33]</td>
</tr>
<tr>
<td>Sevy et al., 2007a</td>
<td>-0.44 [-1.15, 0.27]</td>
</tr>
<tr>
<td>Sevy et al., 2007b</td>
<td>-0.47 [-1.16, 0.23]</td>
</tr>
<tr>
<td>Nakamura et al., 2008</td>
<td>-0.49 [-1.06, 0.08]</td>
</tr>
<tr>
<td>Premkumar et al., 2008</td>
<td>-1.72 [-2.23, -1.21]</td>
</tr>
<tr>
<td>Kim et al., 2009</td>
<td>-0.76 [-1.15, -0.37]</td>
</tr>
<tr>
<td>Lee et al., 2009</td>
<td>-0.95 [-1.43, -0.46]</td>
</tr>
<tr>
<td>Raffard et al., 2011a</td>
<td>-1.39 [-1.89, -0.90]</td>
</tr>
<tr>
<td>Raffard et al., 2011b</td>
<td>-0.52 [-0.98, -0.06]</td>
</tr>
<tr>
<td>Cella et al., 2012</td>
<td>-1.86 [-2.53, -1.19]</td>
</tr>
<tr>
<td>Kim et al., 2012</td>
<td>-0.68 [-1.19, -0.18]</td>
</tr>
<tr>
<td>Fond et al., 2013</td>
<td>-0.71 [-1.07, -0.36]</td>
</tr>
<tr>
<td>Hori et al., 2014</td>
<td>-5.14 [-5.84, -4.44]</td>
</tr>
<tr>
<td>Nestor et al., 2014</td>
<td>-0.38 [-0.73, -0.04]</td>
</tr>
<tr>
<td>Matsuzawa et al., 2015</td>
<td>-0.24 [-0.62, 0.13]</td>
</tr>
<tr>
<td>Stratta et al., 2015</td>
<td>-1.00 [-1.53, -0.47]</td>
</tr>
<tr>
<td>Zegarra-Valdivia, 2015</td>
<td>-2.64 [-3.50, -1.77]</td>
</tr>
<tr>
<td>Zhang et al., 2015</td>
<td>-1.45 [-1.85, -1.04]</td>
</tr>
<tr>
<td>Kim and Kang, 2016</td>
<td>-0.60 [-1.09, -0.12]</td>
</tr>
<tr>
<td>Pedersen et al., 2016</td>
<td>-0.43 [-0.89, 0.02]</td>
</tr>
<tr>
<td>RE Model</td>
<td>-1.06 [-1.50, -0.63]</td>
</tr>
</tbody>
</table>
### Author(s) and Year

<table>
<thead>
<tr>
<th>Author(s) and Year</th>
<th>Standardized Mean Difference [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilder et al., 1998</td>
<td>0.61 [-0.07, 1.29]</td>
</tr>
<tr>
<td>Ritter et al., 2004</td>
<td>0.47 [-0.21, 1.15]</td>
</tr>
<tr>
<td>Rodríguez-Sánchez et al., 2005</td>
<td>-0.04 [-0.51, 0.43]</td>
</tr>
<tr>
<td>Shurman et al., 2005</td>
<td>1.06 [0.33, 1.78]</td>
</tr>
<tr>
<td>Kester et al., 2006</td>
<td>0.66 [-0.00, 1.31]</td>
</tr>
<tr>
<td>Lee et al., 2007</td>
<td>1.45 [0.83, 2.07]</td>
</tr>
<tr>
<td>Martino et al., 2007</td>
<td>0.31 [-0.36, 0.97]</td>
</tr>
<tr>
<td>Sevy et al., 2007a</td>
<td>0.21 [-0.49, 0.91]</td>
</tr>
<tr>
<td>Sevy et al., 2007b</td>
<td>0.10 [-0.58, 0.79]</td>
</tr>
<tr>
<td>Kim et al., 2009</td>
<td>0.65 [0.26, 1.04]</td>
</tr>
<tr>
<td>Kim et al., 2012</td>
<td>0.87 [0.35, 1.38]</td>
</tr>
<tr>
<td>Hori et al., 2014</td>
<td>0.73 [0.38, 1.09]</td>
</tr>
<tr>
<td>Brown et al., 2015</td>
<td>0.21 [-0.18, 0.61]</td>
</tr>
<tr>
<td>Matsuzawa et al., 2015</td>
<td>-0.04 [-0.41, 0.34]</td>
</tr>
<tr>
<td>Zhang et al., 2015</td>
<td>1.26 [0.86, 1.65]</td>
</tr>
<tr>
<td>Kim and Kang, 2016</td>
<td>0.12 [-0.35, 0.59]</td>
</tr>
<tr>
<td>Pedersen et al., 2016</td>
<td>0.16 [-0.29, 0.61]</td>
</tr>
</tbody>
</table>

### RE Model

0.51 [0.29, 0.73]
References


