Dopaminergic Modulation of Probabilistic Reasoning and Overconfidence in Errors: A Double-Blind Study

Christina Andreou*1, Steffen Moritz1, Kristina Veith, Ruth Veckenstedt, and Dieter Naber

Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

*To whom correspondence should be addressed; Martinistrasse 52, 20246 Hamburg, Germany; tel: +49-40-7410-59460, fax: +49-40-7410-59805, e-mail: c.andreou@uke.de

1The authors have contributed equally to the preparation of this manuscript and share first authorship.

Introduction: Reasoning biases such as jumping to conclusions (JTC) and overconfidence in errors have been well replicated in patients with delusions. However, their relation to dopaminergic activity, central to pathophysiologic models of psychosis, has not yet been investigated. This study aimed to examine the effects of a dopaminergic agonist (L-dopa) and a dopaminergic antagonist (haloperidol) on the JTC bias and overconfidence in errors after single-dose administration in healthy individuals. Methods: The study used a randomized, double-blind, placebo-controlled, 3-way crossover design. Participants were 36 healthy individuals aged 18–36 years. The variables of interest were draws to decision and probability threshold to decision on a computerized variant of the beads task and the number of high-confident incorrect responses on a visual memory task. Results: There were no significant effects of substance on draws to decision and probability threshold to decision. A significant effect emerged for high-confident incorrect responses in the memory task; pairwise comparisons indicated a significant reduction of the number of high-confident incorrect responses after administration of haloperidol vs L-dopa and placebo. Conclusions: This is the first study to investigate the direct effects of dopaminergic drugs on reasoning biases. The JTC bias and overconfidence in errors showed a differential pattern of dopaminergic modulation, suggesting that they represent different facets of reasoning abnormalities that interact with each other to produce delusions in susceptible individuals.

Key words: delusions/schizophrenia/reasoning biases/metacognition/jumping to conclusions/liberal acceptance from specific disruptions in the normal cognitive processes governing belief generation and evaluation—the so-called reasoning biases, thinking styles leading to a biased integration and interpretation of perceived stimuli. A central place within this theoretical framework is occupied by the data-gathering bias, a tendency of people with delusions to gather less evidence than controls to arrive at a decision. This abnormality is typically demonstrated using probabilistic reasoning tasks such as the “beads task,” in which participants are presented a number of colored beads (eg, red and blue) in succession and asked to guess which of the 2 jars containing beads in opposite color ratios (eg, 85:15 and 15:85) the presented beads originate from. It has been consistently shown that patients with delusions require fewer draws to reach a decision regarding the origin of the beads compared with healthy controls, ie, they exhibit a jumping-to-conclusions (JTC) bias.

Since the first description of JTC, a number of other reasoning biases associated with delusions have been described. In memory tasks, eg, it has been shown that patients with schizophrenia and delusions are significantly more confident than controls in responses that turn out to be incorrect. This response pattern has been confirmed in domains other than memory, such as decision making and social cognition. Overconfidence in errors has originally been attributed to a “liberal acceptance” reasoning style, ie, a lowered decision threshold for the acceptance of hypotheses. In this aspect, it is conceptually akin to the JTC bias. However, overconfidence in errors additionally entails a dimension of disturbed “evidence integration,” ie, a deficit in the ability to integrate disambiguating or disconfirmatory evidence when the latter does not match the cognitive schema currently held.

The above abnormalities have been well replicated across different populations with delusions and in individuals at high risk or with a genetic predisposition.
for psychosis. However, their pathophysiology still remains unclear. Critically, very little is known about how these reasoning biases relate to the neurobiology of psychotic disorders, particularly with respect to the widely accepted hypothesis of a dopaminergic dysfunction underlying psychotic symptomatology. The latter was originally based on the observation that dopamine (especially D2) receptor antagonists can attenuate psychotic symptoms, while dopamine agonists can aggravate or induce such symptoms. It is postulated, in brief, that the core manifestations of psychosis, such as delusions and hallucinations, are caused by a functional disturbance of the mesolimbic dopaminergic system.

So far, few studies have looked into the relationship between reasoning biases and dopaminergic activity. Most of these studies have looked into this issue indirectly, by investigating the effects of symptom remission after successful antipsychotic treatment on reasoning biases. The JTC bias has been suggested not to change with treatment although it may act as a moderator on treatment response. On the other hand, overconfidence in errors has been consistently reported to be negatively correlated with antipsychotic drug dose, which might imply that it is modulated by dopaminergic activity. Thus, the JTC bias and overconfidence in errors appear to differ in the way they are affected by changes in the dopamine system. However, this interpretation is only a tentative one because only one of the aforementioned studies on the effects of antipsychotics on reasoning biases included patients who were indeed antipsychotic naive at baseline.

A complementary approach would be to investigate the effects of dopaminergic agents on the reasoning biases in question in healthy individuals. This approach carries the additional advantage that it helps shed light on this issue without the confounding influence of acute psychotic symptoms and/or the general intellectual impairments of patients with schizophrenia, who have been the main focus of previous studies. The goal of this study was, therefore, to assess the effects of dopaminergic agonists and antagonists on the JTC bias and overconfidence in errors in healthy participants. The study used a randomized, double-blind, 3-way crossover design to assess the hypothesis that JTC and overconfidence in errors show a differential pattern of dopaminergic manipulation; JTC was assumed not to be influenced by dopaminergic drugs, while overconfidence in errors was expected to increase with dopaminergic agonists and decrease with dopaminergic antagonists. A further aim of the study was to study the effects of dopaminergic agents on a direct measure of liberal acceptance, for which no specific predictions were made, because it has been implicated in both JTC response patterns and overconfidence in errors in previous studies.

### Methods

#### Participants and Design

This study was part of a larger project investigating the effects of dopaminergic agonists and antagonists on cognitive functions associated with psychotic symptoms, such as semantic priming and reasoning biases. Participants were healthy individuals recruited among students and acquaintances of the staff of the Neuropsychology Unit of the Department of Psychiatry, University Medical Center Hamburg-Eppendorf. In order to be included in the study, participants were required to be between 18 and 40 years of age, right handed, and native speakers of German. Exclusion criteria were any past or current psychiatric or neurological disorder (including substance use disorders); a history of schizophrenia or bipolar disorder in a first-degree relative; a history of craniocerebral trauma, arterial hypertension, cardiological conditions, or serious medical conditions; pregnancy; or treatment with any psychotropic or other drugs. Eligibility for the study was confirmed by means of an interview carried out by a trained doctoral-level student. The study was approved by the Ethics Committee of the Medical Association Hamburg and was performed in accordance with the ethical standards laid down in the current version of the Declaration of Helsinki. All participants gave their written informed consent before participating in the study.

In order to assess the effects of dopaminergic agents on reasoning biases, a randomized, double-blind, 3-way crossover design was used. In 3 successive visits, participants were administered either 100 mg l-dopa and 25 mg benserazide (Madopar), 2 mg haloperidol (Haldol), or placebo in randomized order and under double-blind conditions. The dose of l-dopa was identical to that used in previous behavioral and neuroimaging studies. The dose of haloperidol was chosen such as to correspond to a D2 receptor occupancy of around 70%, which is deemed sufficient for a clinical response while minimizing the risk of adverse effects.

The 3 visits were separated by at least 7 days, in order to allow a complete washout of the drug with the longer half-time (haloperidol). In order to compensate for the different $T_{max}$ of haloperidol and l-dopa, a double-dummy design was implemented (table 1). The testing session began thus at the time of maximal serum concentration of each drug and lasted 60 minutes at the maximum. Subjective psychological, somatic, and motor

### Table 1. Double-Dummy Design of Study Drug Administration

<table>
<thead>
<tr>
<th></th>
<th>t0</th>
<th>t1</th>
<th>t2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haloperidol</td>
<td>Haldol</td>
<td>Placebo</td>
<td>Onset of testing session</td>
</tr>
<tr>
<td>Levodopa</td>
<td>Placebo</td>
<td>Madopar</td>
<td>Onset of testing session</td>
</tr>
<tr>
<td>Placebo</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Onset of testing session</td>
</tr>
</tbody>
</table>
(adverse) effects of the drugs were assessed through subjective ratings on a scale using several Likert-type statements at the time of ingestion of the second capsule (t1) and after the end of the testing session; moreover, blood pressure and pulse were measured at 30-minute intervals. In order to ensure blinding, participants were asked to guess which substance they had received at the end of each session.

**Tasks and Procedure**

The following tasks were used for the assessment of reasoning biases:

- A computerized variant of the classical probabilistic reasoning task, in which beads in jars have been replaced by fish in lakes. Participants were shown 2 lakes containing red and blue fish, lake A with 60 red and 40 blue fish and lake B with the reverse ratio. A ratio of 60:40 was used, in order to increase the degree of difficulty of the task and, accordingly, its discriminatory power in healthy participants (cf Evans et al for a relevant discussion). Ten fish were successively presented in a predetermined sequence to the participant; following each draw, participants were asked to indicate whether they had arrived at a decision regarding the origin of the fish (and, if so, which lake it came from) and additionally provide a probability rating as to the possibility that the fish originated from lake A. All fish drawn remained visible throughout the task in order to minimize working memory demands. Moreover, the task was not terminated as soon as the participant reached a decision but after the presentation of the final fish (graded estimates procedure with simulated decision). Thus, it was possible for participants to reconsider their probability ratings and/or decision with each successive fish. This task has been shown to be equally sensitive to the classical beads task for measurement of the JTC bias. The variables of interest were (a) the number of draws to decision because this variable has been shown to be the most adequate measure of the JTC bias in a recent meta-analysis and (b) the probability threshold, at which a decision was made, as a measure of liberal acceptance, one of the suggested underpinnings of JTC.

- A visual variant of the Deese-Roediger-McDermott paradigm was used to assess overconfidence for false memory judgments. In each session, 2 black-and-white, easily identifiable, prototypical pencil-drawn scenes (eg, classroom and beach) were presented to participants for 40 seconds each. In the ensuing recognition trial 10 minutes later, a total of 24 previously presented items and 24 new (distractor) items (12 each for every scene) were presented verbally on the screen along with a contextual cue indicating the scene they referred to (classroom). The distractor items included previously not presented items that were either unrelated to the scene in question (n = 4 for each scene) or related to some central aspect of the scene (eg, towels for the beach scene, n = 8). Participants were required to indicate whether they recognized the presented item and their degree of confidence in their response. The variables of interest were the number of high-confident incorrect responses, which has been shown to successfully discriminate between patients with schizophrenia and healthy controls.

The tasks were always administered in the same order (fish task first). Both tasks were available in 3 parallel versions, in order to minimize practice effects. Regarding the fish task, there was concern that different sequences might induce high variability in response patterns. Thus, the 3 versions differed in the color of the fish and the lake they were drawn from (A or B) but not in the sequence used, similar to previous studies in patients with schizophrenia. For the false memory task, different scenes were used in each session. The various versions of each task were presented in a fixed order across visits, while the order of drug administration was randomized. In this way, performance measures on each substance relied on data from all 3 parallel versions, thereby minimizing version-specific effects.

In order to rule out performance differences due to nonspecific effects of the drugs on attention, the d2 test was also administered at each session. The d2 test is a letter cancellation test consisting of the letters d and p, arranged in 14 rows of 47 letters each. The letters are marked with 1–4 dashes, and participants are requested to cross out only the letter d with 2 dashes (both above, both below, or above and below the letter). Twenty seconds are allowed for each row. The total score consisted in the number of correctly processed items minus incorrectly crossed out distractors. The test has well-documented validity and an excellent test-retest reliability.

Participants were paid a total of 80€ (or 40€ plus course credit for students) for their participation in the study or a proportional amount in case of dropout before all 3 sessions were completed.

**Statistical Analyses**

Repeated-measures ANOVAs with substance (haloperidol vs l-dopa vs placebo) as the within-subjects factor were carried out on JTC, liberal acceptance and error overconfidence measures, and d2 scores.

In the false memory task, the number of “old” responses (ie, items recognized as previously presented) was analyzed per substance (l-dopa, haloperidol, and placebo) and item type (previously presented vs distractor items). This analysis was conducted as a validity check, lacking a direct measurement of drug serum levels. In a recent study, l-dopa has been shown to increase false positive responses compared with dopaminergic antagonists in a memory paradigm, while leaving overall memory...
performance unaffected.\textsuperscript{43} Emergence of a similar pattern in this study would constitute indirect proof that the administered drugs produced the intended pharmacologic effects.

Differences between the 3 substances regarding adverse effects were assessed by means of a 3 (substance) \times 3 (time: t0, t1, and t2) repeated-measures ANOVA. Finally, in order to confirm blinding, a Pearson's chi-square test was conducted between ingested substance and substance guessed by the participants at the end of the session.

Results

Thirty-six participants (20 males; mean age 24.3 ± 4.0, range 19–36) completed all 3 testing sessions. Of these, 1 participant was excluded from overconfidence analyses due to extreme outlier responses in the false memory paradigm. Mean scores per substance are presented on Table 2.

There were no significant differences among the 3 substances in d2 scores (F(2,70) = 1.28, \(P = .3\)), nor in adverse effects (main effect of substance: \(F(2,70) = 0.82, \ P = .4\); time \times \text{substance interaction}: \(F(4,140) = 1.62, \ P = .2\)). There were no dropouts and no premature session terminations due to adverse effects. There was also no association between ingested and guessed substance (\(\chi^{2}(6) = 5.03, \ P = .5\)).

In the task, substance did not have a significant effect on either draws to decision (F(2,70) = 0.04, \(P = .9\), \(\eta^{2} = 0.002\)) or decision threshold (F(2,70) = 0.93, \(P = .4\), \(\eta^{2} = 0.026\)). In order to exclude that this lack of significance was due to practice effects, a follow-up ANOVA was conducted on data from the first testing session only; again, the effect of substance was not significant (draws to decision: F(2,33) = 0.54, \(P = .6\); decision threshold: F(2,33) = 0.71, \(P = .5\)).

Regarding confidence, there was a significant effect of substance on the number of high-confident incorrect responses (F(2,68) = 3.10, \(P = .05\), \(\eta^{2} = 0.08\)). Post hoc pairwise comparisons (Fisher’s least significant difference method) showed significantly less high-confident incorrect responses with haloperidol compared with 1-dopa (\(P = .03\)) and placebo (\(P = .03\)). The total number of high-confident responses (correct and incorrect), in contrast, did not differ among substances (F(2,68) = 1.12, \(P = .3\)).

As expected, memory accuracy was not affected by substance (F(2,68) = 2.80, \(P = .10\)), but there was a significant substance \times item-type interaction (F(2,68) = 5.00, \(P = .01\)). Post hoc simple contrasts indicated that this reflected a higher rate of false alarms (new items recognized as previously presented ones) for 1-dopa compared with haloperidol (F(1,34) = 8.57, \(P = .006\)).

Discussion

This study used a randomized, double-blind, 3-way crossover design to investigate the effects of dopaminergic agonists and antagonists on reasoning biases associated with delusions. The main findings were that the JTC bias and liberal acceptance (reflected, respectively, in draws to decision and probability threshold to decision in the JTC task) did not appear to be affected by dopaminergic agents, whereas there was significantly decreased overconfidence in memory errors with haloperidol compared with 1-dopa and placebo.

Our results regarding the JTC bias and liberal acceptance are in accordance with previous findings from studies in clinical populations, where successful treatment with antipsychotics has shown not to affect data-gathering abnormalities.\textsuperscript{14,23,44} Nevertheless, the latter are robustly associated with the presence or propensity for delusions (see “Introduction” section). Thus, the question arises how exactly these reasoning biases fit into the dopamine hypothesis of psychosis. In this context, 2 theoretical accounts of delusions are relevant. First, it has been convincingly argued that dysregulated dopaminergic activity leads to aberrant allocation of salience to random stimuli, such that they are overly weighted.\textsuperscript{45–47} Second, it has been proposed\textsuperscript{1} that delusions result from 2 distinct types of cognitive disturbance, which occur in combination: First, an implausible thought is generated; second, the thought is accepted uncritically as true. It could be assumed that a state of heightened salience of stimuli corresponds to the

<table>
<thead>
<tr>
<th>Table 2. Results of Cognitive Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Jumping-to-conclusions task</td>
</tr>
<tr>
<td>Draws to decision</td>
</tr>
<tr>
<td>Probability threshold</td>
</tr>
<tr>
<td>False memory task</td>
</tr>
<tr>
<td>Correctly recognized items</td>
</tr>
<tr>
<td>False positive errors</td>
</tr>
<tr>
<td>High-confident responses</td>
</tr>
<tr>
<td>High-confident errors</td>
</tr>
<tr>
<td>d2 test</td>
</tr>
</tbody>
</table>
first type of disturbance, eg, by rendering neutral events relevant for the self, whereas impaired probability reasoning may come into play later, leading to premature acceptance of an implausible thought. 13 This assumption is consistent with findings of a previous study, 22 in which the JTC bias in patients with schizophrenia was indeed associated with a lowered threshold for responding but not with increased propensity to overweight evidence. It also corroborates the suggestion by Menon et al 14 that antipsychotics reduce delusional intensity by decreasing aberrant salience rather than by influencing the JTC bias per se.

It might be argued that the failure to obtain significant results in the JTC task was due to a lack of statistical power or to the conceptual simplicity of the fish task, leading to practice effects despite the use of parallel versions. In our view, these interpretations are not very plausible. The observed effect sizes were very weak, thus excluding the possibility that a larger sample would lead to different findings; a post hoc calculation also showed that the sample size necessary to demonstrate an effect of at least the same size as the one obtained for overconfidence ($\eta^2 = 0.08$) with a power of 0.8 would be 28, i.e., smaller than the one used in this study. 48 Moreover, a follow-up analysis including only data from the first session (i.e., before the emergence of practice effects) did not produce any significant differences among the 3 substances. Finally, another recent study 36 has also not observed any significant modulation of response patterns on the JTC task using ketamine, another drug associated to the emergence of psychotic symptoms and, indirectly, 49 to the dopaminergic system.

Quite in contrast to the JTC bias, overconfidence in errors did show a pattern of dopaminergic modulation. To our knowledge, this is the first direct evidence of an association between a delusion-associated reasoning bias and the dopaminergic system. The finding of reduced overconfidence in errors with haloperidol is consistent with previous reports on the effect of antipsychotics on error overconfidence in clinical populations. 24, 25 Interestingly, another reasoning bias conceptually related to error overconfidence and strongly associated with delusions in the literature, 50–53 belief inflexibility, has also been suggested to improve with antipsychotic medication. 23

It appears thus that although data-gathering biases are not linked with dopaminergic processes in themselves, they might interact with other cognitive abnormalities resulting from dopaminergic dysregulation such as aberrant salience or “incorrigibility” (an umbrella term used henceforth to subsume the concepts of overconfidence in errors and belief inflexibility), leading to the emergence of the clinical syndrome of delusions. According to existing accounts, salience exerts its effects at an early stage of delusion formation, providing the ground for the generation of delusional ideas; in contrast, it has been suggested 51 that belief inflexibility largely mediates the contribution of a JTC reasoning style to delusional conviction, thus playing a role at a later stage of delusion consolidation. However, the hypothesized association of both salience and incorrigibility with dopaminergic activity raises the possibility that these disturbances reflect 2 facets of the same core deficit, interacting with the JTC bias to produce delusions. As to what this deficit might be, only speculations can be made at this point. For example, borrowing from a prevalent model of schizophrenia, 34–36 aberrant salience and incorrigibility could both result from a deficit in establishing and updating, respectively, contextual information (see also Langdon et al 11 p. 328, for a related account of delusion formation). Alternatively, excessive salience and incorrigibility could be conceptualized as resulting from disturbed processing of feedback information within the reward system because the processing of previous outcomes modulates the attentional bias toward stimuli (incentive salience). 57 Both accounts are neurophysiologically plausible because dopamine is implicated both in working memory “gating” that enables flexible context updating and interference protection 58 and in the processing of prediction error. 59 However, these hypotheses are only tentative and need to be directly tested in future studies.

Certain limitations of the study need to be acknowledged. Despite the fact that memory performance differences between haloperidol and l-dopa were, as an indirect validity check, in the expected direction, 45 drug plasma-level measurements would have substantially added to the validity of our findings—especially because the doses used were in the low range for both agents. In the case of haloperidol, the issue is probably of less importance. Although haloperidol exhibits substantial interindividual variability in its pharmacokinetics, 90 the relationship between oral dose and receptor occupancy has been well documented, at least in the steady state, 34 and the observed effects in this study were quite robust. The absence of plasma-level measurements is more relevant in the case of l-dopa because there were no differences between l-dopa and placebo in the false memory task in this study. It is possible that the dose of l-dopa might have been too low to detect significant differences from placebo in the particular paradigm used in this study because similar doses of l-dopa have been shown to successfully modulate various aspects of cognition in previous studies. 27, 30, 32, 43, 61 To the best of our knowledge, no studies have investigated the dose-effect relationship of l-dopa on cognitive performance so far, which might be an interesting goal for future studies.

Another limitation of the study results from its design, which consisted in single-dose administration of dopaminergic agents to healthy subjects, thus limiting generalizability of findings to clinical populations. As delineated in the introduction in more detail, the choice of this design is advantageous in that it circumvents some of the problems associated with the study of clinical populations. Studies using similar doses of l-dopa in healthy volunteers have been used to draw inferences concerning the generation of psychotic symptoms. 27, 43 Furthermore, single-dose haloperidol has also been used successfully to model cognitive adverse effects...
of antipsychotic treatment\textsuperscript{62} and has been reported to reduce cannabis-induced psychotic symptoms in healthy individuals to an extent comparable to clinically relevant reductions in patient populations\textsuperscript{63}; moreover, it has been shown that the relationship between dopamine antagonist dose and receptor occupancy is similar in healthy individuals as in chronically treated patients with schizophrenia.\textsuperscript{64} However, psychotic disorders and their treatment are associated with significant alterations in several aspects of dopaminergic system function\textsuperscript{65,66} which are probably approximated only to a limited extent by single-dose administration of dopaminergic agents to healthy individuals. Moreover, no measures of delusional propensity or delusion-like experiences were included in the assessments. Hence, any attempts to extrapolate our results to the clinical syndrome of delusions should be undertaken with caution and only in consideration of findings from studies investigating clinical populations.

In summary, the JTC and liberal acceptance biases were not influenced by the administration of dopaminergic agonists and antagonists. In contrast, dopaminergic antagonists tended to reduce overconfidence in errors. The results suggest that JTC and overconfidence in errors represent different facets of abnormalities that interact with each other to produce delusions in susceptible individuals. Further studies are warranted to confirm the generalizability of findings in clinical populations.

**Funding**

German Research Foundation, Deutsche Forschungsgemeinschaft (AN970/1-1 to C.A.).

**Acknowledgments**

The authors would like to thank Anna Schildt and Aljoscha Rieger for their help with recruitment and data acquisition, as well as 5 anonymous reviewers for their constructive remarks and helpful suggestions. The authors have declared that there are no conflicts of interest relating to the subject of this study.

**References**


49. Kapur S, Seeman P. NMDA receptor antagonists ketamine and PCP have direct effects on the dopamine D(2) and serotonin 5-HT(2) receptors-implications for models of schizophrenia. Mol Psychiatry. 2002;7:837–844.


