**Bayesian Accounts and Black Swans**

**Questioning the Erotetic Theory of Delusional Thinking**

Ryan McKay

ARC Centre of Excellence in Cognition and its Disorders,

Department of Psychology, Royal Holloway, University of London

**Acknowledgements:** I thank Rob Ross and Kengo Miyazono for valuable comments on a draft of this manuscript.

**Address for correspondence:** Ryan McKay, Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom. Tel: +44 (0) 1784 41 4993; Email: [ryantmckay@mac.com](mailto:ryantmckay@mac.com)

**Abstract**

Parrott and Koralus argue that a particular cognitive factor – “impaired endogenous question raising” – offers a parsimonious account of three delusion-related phenomena: 1) the development of the Capgras delusion; 2) evidence that patients with schizophrenia outperform healthy control participants on a conditional reasoning task; and 3) evidence that deluded individuals “jump to conclusions”. In this response I assess these claims, and raise my own questions about the “erotetic” theory of delusional thinking.

**Keywords:** Delusions**;** Erotetic Theory;Jumping to Conclusions; Schizophrenia.

Parrott and Koralus (P&K; this issue) present a fresh take on an important topic for cognitive neuropsychiatry.They argue that a particular cognitive factor – “impaired endogenous question raising” – offers a parsimonious account of three delusion-related phenomena: the development of the Capgras delusion, evidence that patients with schizophrenia outperform healthy control participants on a conditional reasoning task, and evidence that delusional patients “jump to conclusions”. P&K suggest their account is superior to “Bayesian” alternatives, which lack simplicity and unity.

In this response to their paper I start out by examining the core claim that deluded individuals have selectively impaired endogenous-question-raising. I then assess P&K’s specific erotetic accounts of the three aforementioned phenomena, before contesting the claims that the erotetic theory is more parsimonious and unified than Bayesian alternatives. I conclude with a summary of my questions for P&K.

**Do deluded individuals have a selective deficit in endogenous-question-raising?**

P&K’s theory hinges on a key distinction between self-initiated (endogenous) and externally stimulated (exogenous) questions; they maintain that deluded individuals are selectively deficient in raising the former sort. This should mean a) that the ability of deluded individuals to raise endogenous questions is impaired; and b) that their response to exogenous questions is intact. However, both of these claims seem problematic to me.

The first claim seems problematic because it is not clear what counts as an endogenous question. The provenance of some questions is obvious – in particular, questions explicitly posed by other agents are unambiguously exogenous – but P&K distinguish “default questions in response to external stimuli” (which they also want to classify as exogenous) from endogenous “follow-up questions”. For example, they “suggest that generally if somebody appears in front of us, regardless of whether we are delusional, this naturally raises the question of who this person is among people we know.” According to P&K, this question is such an obvious “default” response to this external stimulus that it cannot be viewed as self-initiated. But P&K do not provide any rigorous way of distinguishing a “self-initiated” question from an “externally stimulated” question, which makes it difficult to test (and potentially falsify) their claim that deluded individuals are impaired in raising the former. I think that the erotetic theory would benefit from a precise specification of the endogenous-exogenous distinction.

The second problem concerns the claim that the response of deluded individuals to exogenous questions is intact. P&K say that someone with a deficit in endogenous-question-raising “would have no problem taking on board and answering questions that are put to her by someone else”, but by their own admission people with delusions *don’t* take such (unambiguously exogenous) questions properly on board (P&K talk of the rubber band that snaps back to its original position once external prompting is removed). As far as I can see, P&K’s theory does not offer a compelling account of why the rubber band snaps back – if question *raising* is the problem, they need to explain why deluded individuals cannot properly utilize and retain questions raised by others (more on this below).

**Does the erotetic theory shed light on Capgras delusion?**

P&K’s explanation of Capgras delusion (the belief that a loved one has been replaced by an impostor) proceeds as follows:

* We represent other people we know as bundles of features. For example, I might represent my friend Jack as having blonde hair, being tall, and speaking with a South African accent.
* We represent our relationships with other people as features of those other people. If Jack is a particularly *close* friend of mine, the bundle of features comprising my representation of him will include (in addition to his hair colour, height and accent) what P&K call the “C-feature.”
* When presented with a person walking into a room, the question “who, if anyone, is this among people I know?” automatically occurs to us (according to P&K, we don’t initiate it ourselves; rather, the question is “externally stimulated” by that person’s appearance).
* If the person who has just entered the room has a feature that characterizes none of the people we know, or lacks a feature that characterizes everyone we know, the answer we compute for “who is this among people I know?” will be “nobody I know.” If the person is my friend Jack, and if, owing to a neurophysiological impairment, I represent him as lacking the C-feature, I will fail to identify him as Jack.
* [Here I quote P&K directly] “If we hold this conclusion fixed, it does not seem like a further mistake on the part of the patient to speculate on various impostor scenarios...”

As far as I can see, the last point does not follow from those that precede it: It *does* seem to me a mistake for the person to speculate on various impostor scenarios here. P&K do not imply that the C-feature has any special status (they say “Someone might represent his wife as having the C-feature in the same way in which he might represent her as having a certain eye color”). So at this stage all I know is that a person with some, but not all, of Jack’s features has entered the room. Why would I speculate that the person must be someone posing as my friend Jack? Why not just conclude that this person is a stranger? After all, there must be thousands of tall blonde South Africans who are strangers to me. P&K give no answer to this. Moreover, their suggestion that deluded individuals in this situation would “speculate on various impostor scenarios” seems inconsistent with the core tenet of their theory: that deluded individuals fail to engage in the sorts of endogenous reflection that healthy individuals do. P&K might respond that impostor speculations are a “default” response in this situation, but their analysis gives no reason to think that (again, the obvious default is just “this person is a stranger”).

A related problem is that if the absence of the C-feature prompts impostor speculations when I encounter Jack, then the absence of any other Jack-feature should do so as well. If Jack dyes his blonde hair brown, then I should immediately speculate that he is an impostor; and if I have impaired endogenous-question-raising, I should fail to revise this speculation, adopting it as a belief. So why is Capgras delusion not more common? It seems to me that on P&K’s analysis, Capgras delusion should develop every time the acquaintances of a person with impaired endogenous-question-raising change their appearance (provided that the relevant aspect of appearance is one of the features in the representation bundle).

A further issue concerns P&K’s claim that the response of deluded individuals to exogenous questions is intact. Although some deluded patients are indeed responsive to external questioning, others are not:

MF’s indifference to constant questioning about his delusion was striking. He was unperturbed when evidence contradicting his delusion was pointed out to him, and when the examiner repeatedly emphasised the improbability of the delusion. MF understood that the research the examiner was conducting involved investigating delusions, and he understood that the particular delusion being investigated was his belief that the woman living with him was not his wife. (Breen, Caine & Coltheart, 2002, p. 120)

In this case, the prospect that the patient was deluded was explicitly raised through exogenous questioning, yet the delusion persisted even in the face of those questions.

In sum, it is unclear to me how the notion that deluded individuals have selectively impaired endogenous-question-raising sheds light on the Capgras delusion. I do not see a) why representing someone as lacking the C-feature would prompt the Capgras delusion in a person with deficient endogenous-question-raising, and b) why the delusion would persist if the person’s response to exogenous questions were intact.

**Does the erotetic theory illuminate conditional reasoning in Schizophrenia?**

P&K next turn their attention to an interesting study by Mellet et al. (2006), who found that patients with schizophrenia outperformed healthy control participants on a particular reasoning task. The task involved falsifying conditional rules with a negation in the antecedent, such as: *If there is not a red square on the left, then there is a yellow circle on the right*.

P&K suggest that healthy participants adopt what they call a “suppositional reasoning strategy” to solve such problems. As I understand their argument, this strategy involves finding a state of affairs that confirms the affirmative clause and contradicts the negative clause of the rule in question (e.g., the state of affairs *a yellow circle on the right* confirms *there is a yellow circle on the right;* the state of affairs *a red square on the left* contradicts *there is not a red square on the left*). For conditional rules with a negation in the consequent (e.g., *If there is a red square on the left, then there is not a yellow circle on the right*), the suppositional strategy amounts to finding a state of affairs that confirms the antecedent and contradicts the consequent, which is the correct logical strategy. But for conditional rules with a negation in the antecedent, the suppositional strategy amounts to finding a state of affairs that contradicts the antecedent, which is an *in*correct logical strategy.

P&K are right that the “suppositional” strategy will yield correct answers for cases with a negation in the consequent, and incorrect answers for cases with a negation in the antecedent. So I agree that the results of Mellet et al. (2006) are consistent with healthy participants (tending to) adopt this strategy and patients with schizophrenia (tending to) adopt the correct strategy. What I find unclear is why the suppositional strategy should require a capacity for endogenous-question-raising, whereas the correct logical strategy shouldn’t. Indeed, P&K say “[W]e have good reasons to think that individuals often use suppositional reasoning to address problems that otherwise require us to represent many alternative possibilities”, which suggests the suppositional strategy would be perfect for people with *deficient* endogenous-question-raising (because elsewhere P&K argue that such people favour questions that represent fewer alternatives; more on this below).

In sum, it seems to me that P&K do not offer a compelling reason why deficient endogenous-question-raising would give rise to the pattern of responses produced by the patients with schizophrenia in Mellet et al’s (2006) study. However, whether this is a significant problem for P&K’s theory depends on the scope they intend it to have. Insofar as their theory is an account of delusional cognition – which is what the title and abstract of their paper suggest – then there is no major problem, because delusions (a symptom) and schizophrenia (a diagnostic syndrome) are not the same thing. Not all people with delusions have a diagnosis of schizophrenia, and not all people with a diagnosis of schizophrenia have delusions. Advocates of a symptom focused approach (which is arguably the core approach of cognitive neuropsychiatry; see Langdon, 2011[[1]](#endnote-1)) note that the individual symptoms of syndromes like schizophrenia sometimes have divergent associations with specific patterns of responses. For example, Bentall, Baker and Havers (1991) found that hallucinating patients were more likely to misattribute self-generated events to an external source than were patients with delusions but no history of hallucinations. Although P&K say that “a delusion-related diagnosis of schizophrenia can in fact improve performance on reasoning tasks”, there is no evidence that the improved reasoning of the patients in Mellet et al’s (2006) study was related to delusions specifically. So P&K have no need to explain this finding if their intention is to explain delusional cognition.

**Does the erotetic theory explain the tendency of deluded individuals to “Jump to Conclusions”?**

That delusional and delusion-prone individuals “jump to conclusions” on the well known “beads task” is one of the most important and influential findings in the literature on delusions (Huq, Garety & Hemsley, 1988; Fine, Gardner, Craigie & Gold, 2007; Garety and Freeman, 2013). The task of participants in the beads task is to determine from which of two jars a sequence of coloured beads is being drawn, and to indicate when they are ready to make this determination. One of the jars contains beads of mostly one colour, while the other jar mostly contains beads of a contrasting colour (e.g., Jar A may contain 85 red beads and 15 blue, while Jar B contains 85 blue beads and 15 red). A robust finding is that participants who are more prone to delusions require fewer draws before deciding which of the two jars the beads are being drawn from.

How might P&K’s theory account for this finding? In “erotetic” terms, participants face two relevant questions on each trial of the beads task: Q1) Are the beads being drawn from Jar A or Jar B? and Q2) Should I request more information before answering Q1? Could it be the case, therefore, that participants with deficient endogenous-question-raising simply neglect to ask themselves the second of these questions?

The problem with this suggestion is that participants in the beads task don’t *need* to ask themselves these questions, because these questions are explicitly asked by the experimenter on each trial. Perhaps aware of this, P&K try a different tack. According to P&K, participants answer the question of whether the beads are being drawn from Jar A or Jar B by contemplating the pattern of beads in the sequence. Certain patterns are “diagnostic of the bias inherent in Jar A” (e.g., R-R-R-B-B-R-R-R) while others suggest Jar B. This seems uncontroversial. What P&K then suggest, however, is that individuals with a reduced tendency to endogenously raise questions will ask themselves questions that explicitly represent fewer bead sequences, the minimal such question being “do I have Jar A, which has red on its first draw, or do I have Jar B, which has blue on its first draw?”

I’m not sure I follow the logic of this claim, as I don’t see the equivalence between raising fewer questions and raising “lesser” questions (questions that represent fewer alternatives). P&K claim these are equivalent, but they don’t explain why. In any event, without a clear distinction between endogenous and exogenous questions, the erotetic approach seemingly amounts to suggesting that deluded individuals are less inquisitive than healthy individuals. I’m not sure how far this goes beyond previous theoretical proposals[[2]](#endnote-2), or how much it adds to what we already know from dozens of studies on the jumping-to-conclusions bias, which is that deluded individuals seek less evidence than healthy individuals when forming beliefs and making decisions.

**Is the erotetic theory more parsimonious than “Bayesian” alternatives?**

P&K repeatedly claim a key virtue of their account is that it does not posit or appeal to “intrinsically irrational” states or processes. For example, they say:

[N]othing about the proposed cognitive *processes* is intrinsically irrational or intrinsically different from those we would find in ordinary individuals. What differs is merely the extent to which certain processes (e.g. self-generated question-raising) are available.

I take P&K to be implying that whereas some theories posit a qualitative difference between deluded and healthy individuals, the erotetic theory only posits a quantitative difference. There is some precedent for the view that such “qualitative parsimony” is a virtue (Lewis, 1973).[[3]](#endnote-3) However, what I find confusing is that at other points P&K appear themselves to endorse a qualitative difference between deluded and healthy individuals:[[4]](#endnote-4)

The difference between delusional and normal thinking is located in the lack of a safeguard against drawing misguided conclusions, namely a lack of a tendency to endogenously raise further questions.

Here P&K construe deluded individuals as having a cognitive deficit: they *lack* something that healthy individuals have (cf. Langdon & Coltheart, 2000[[5]](#endnote-5)). In contrast, some “Bayesian” accounts preserve the qualitative parsimony that is inherently missing in deficit models (for what that is worth). For example, the notion that deluded individuals overweight current evidence when updating beliefs (McKay, 2012;[[6]](#endnote-6) see also Adams, Stephan, Brown, Frith & Friston, 2013;[[7]](#endnote-7) Menon, Pomarol-Clotet, McKenna & McCarthy, 2006; Speechley, Whitman & Woodward, 2010) does not entail that healthy individuals don’t *also* do this (albeit to a lesser extent; Kahneman & Tversky, 1973). Indeed, recent beads task evidence that P&K outline is consistent with the idea that whereas all individuals overestimate the diagnosticity of current evidence, in high-delusion-prone participants this tendency is exaggerated. I will briefly unpack this below.

In the beads task the colour of a drawn bead is diagnostic of the corresponding jar. For example, a red bead is diagnostic of the mostly red jar, because red beads are what one would tend to expect if the experimenter were drawing from that jar:

***Evidence***: “The current bead is red”

*Is diagnostic of*

***Scenario***: “The experimenter is drawing beads from the mostly red jar”

To the extent that a participant in the beads task overweights the diagnosticity of current evidence, the colour match between the current bead and the corresponding jar will exert greater influence on belief and decision than it should do, and the participant will be more inclined to pick the jar in question at this point than he or she should be (“jumping to conclusions”). Using an incentivized beads task, Furl and Averbeck (2011) found that healthy participants requested fewer pieces of evidence than a Bayesian observer would have done. This finding was replicated by Van der Leer et al. (2015), who also found that this tendency was greater the more prone to delusions the participants were.

The notion that deluded individuals overweight diagnosticity when updating beliefs is parsimonious not just in the sense that it need not posit a qualitative difference between deluded and healthy individuals. More importantly, this idea offers a neat explanation of a number of specific delusions, including the Capgras delusion. As P&K note, Capgras patients are thought to have deficient autonomic responses to familiar faces. When such individuals encounter a loved one, they may be confronted with the following evidence (I leave aside for now the question of whether this evidence is consciously experienced):

***Evidence***: “That person looks just like my loved one but something is not right”

The hypothesis that the loved one has been replaced by an impostor (though at odds with many background beliefs) matches this evidence quite nicely. In other words, this evidence is diagnostic of the impostor scenario – the evidence may be just what one would expect if that scenario obtained (see Coltheart, Menzies & Sutton, 2010):

***Evidence***: “That person looks just like my loved one but something is not right”

*Is diagnostic of*

***Scenario***: “My loved one has been replaced by an impostor”[[8]](#endnote-8)

To the extent that the individual overweights diagnosticity and underweights considerations of general plausibility, the match between the evidence and the impostor hypothesis will exert more influence on posterior belief than it should, perhaps enough to produce the Capgras delusion.

**Is the erotetic theory more unified than “Bayesian” alternatives?**

According to P&K, their erotetic theory offers a simpler, more unified explanation than Bayesian models. In light of the above discussion, it is not obvious to me that P&K’s account is more parsimonious than existing alternatives: indeed, it is possible to argue the opposite. But what about theoretical unity? Here P&K argue as follows:

We think that a theoretically informed model should strive to explain a range of data in a fairly unified manner. A potentially problematic aspect of the Bayesian account seems to be that the difference between control and delusion participants is, in certain cases like the Capgras delusion, explained by a failure of Bayesian rationality, but, at least in the classical version of the Beads Task, it is explained by particularly strong Bayesian rationality. Moreover, in the incentivized version of the Beads Task in which both low and high delusion-prone subjects depart from a Bayesian ideal, it is not clear how to give a Bayesian explanation of this difference. Since both groups ‘jump to conclusions’ more than would prescribed by a normative Bayesian model, in what sense is their reasoning Bayesian? In each case, the relevant Bayesian “explanations” seem more like formal illustrations of the phenomena in question than explanations of a particular cognitive factor that may be implicated in delusional cognition.

In my view, the apparent disunity of “Bayesian” accounts is largely due to P&K’s classification scheme. In particular, they treat as “Bayesian” any account that suggests delusions involve a *deviation* from Bayesian reasoning (in a footnote, P&K raise the issue of whether such “deviation” accounts are really Bayesian, but they opt to set this question aside). Their approach, therefore, involves lumping together theories that argue *p* (delusions are formed in accordance with Bayesian reasoning) with theories that argue *not-p* (delusions are *not* formed in accordance with Bayesian reasoning) and then criticizing the resulting collection as lacking unity.

More worryingly, when some behavior of deluded individuals seems incompatible with strictly Bayesian reasoning, P&K view this as a problem for “Bayesian” approaches, seeming to forget that their “Bayesian” category includes deviation-from-Bayesian-reasoning theories. In the quote above, for example, they say it is not clear how to give a Bayesian explanation of the Van der Leer et al. (2015) findings: “Since both groups ‘jump to conclusions’ more than would prescribed by a normative Bayesian model, in what sense is their reasoning Bayesian?” One response is that the reasoning of high delusion-prone participants is *not* Bayesian – it *deviates* from Bayesian reasoning.[[9]](#endnote-9) But according to P&K, this is itself a Bayesian response. To my mind, P&K should reconsider either their definition of “Bayesian” accounts or their claim that Bayesian accounts cannot account for this finding.

P&K’s treatment of Capgras delusion provides another example of this problematic approach. They say, “a Bayesian account of the Capgras delusion presupposes either some intrinsically irrational prior probability distribution or some kind of probabilistic bias. But, there is no clear way to give a Bayesian explanation of either of these.” On the face of it, this is correct – Bayes theorem says nothing about the prior distribution one should adopt, and a probabilistic bias is inherently non-Bayesian (it’s a bias precisely because it’s a deviation from Bayesian reasoning). But again, on P&K’s definition of “Bayesian” the “deviation from Bayesian” account is itself Bayesian, so it is inconsistent to say there “is no clear way to give a Bayesian explanation” of this.

To be clear, I have no problem with using the term “Bayesian accounts” to refer to accounts that use Bayesian reasoning as a general framework (even if some such accounts posit departures from this framework). The problem is that P&K use the term both in this general way and in a more specific way, to denote accounts that claim delusional reasoning is approximately Bayesian. As it stands, their approach to “Bayesian” accounts seems akin to the following strategy:

1. Collect together the claims “all swans are white” and “not all swans are white”.
2. Call the resultant collection “white swan accounts”.
3. Criticize the disunity of such accounts.
4. Argue that evidence of black swans is a problem for such accounts.

In my view this compromises their attempt to show that their own model is an improvement over the Bayesian model.

**Conclusion**

In the spirit of the erotetic theory, I close with a recap of some questions P&K might want to consider as they further develop their model:

* How can we reliably distinguish “externally stimulated” questions from “self-initiated” questions? In particular, how can we distinguish “default questions in response to external stimuli” from endogenous “follow-up questions”?
* If deluded individuals are selectively deficient in *raising* questions, why are they unable to fully utilize and retain questions that others raise? Why are some deluded individuals completely impervious to exogenous questions? And why does the rubber band snap back for those who are initially responsive to such questions?
* If I encounter a person lacking the C-feature, what could lead me to speculate that the person is an impostor – that they are *posing* as somebody I know? Would the absence of the C-feature need to be accompanied by the presence of some minimal subset of the other features in the representation bundle of a person I know?
* If the absence of the C-feature prompts impostor speculations when I encounter someone I know, why would the absence of other relevant features not do the same? If I have deficient endogenous-question-raising and I represent my doctor as wearing a white coat, why would I not develop Capgras delusion if I see her without the coat?
* What is the intended scope of P&K’s theory? Is it a theory of delusions, of schizophrenia, or of both delusions and schizophrenia?
* What is so “suppositional” about the suppositional reasoning strategy? Why would this strategy require a capacity for endogenous-question-raising, whereas the correct logical strategy wouldn’t?
* Is there a potential dissociation between raising fewer questions and raising “lesser” questions (questions that represent fewer alternatives)?
* Does the suggestion that deluded individuals are less inquisitive than healthy individuals add much to what we already know from studies of data gathering in deluded individuals?

The erotetic theory is certainly intriguing, but some of the details seem elusive at this point. I feel that answers to the above questions would enable a fuller assessment of this interesting theory’s contribution to understanding delusional cognition.

**References**

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D. & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry, 4*, 47.

Aimola Davies, A. M. & Davies, M. (2009). Explaining pathologies of belief. In M. R. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience: philosophical perspectives* (pp. 285-323).

Bentall, R. P., Baker, G. A. & Havers, S. (1991). Reality monitoring and psychotic hallucinations. *British Journal of Clinical Psychology, 30*, 213–222.

Breen, N., Caine, D. & Coltheart, M. (2002). The role of affect and reasoning in a patient with a delusion of misidentification. *Cognitive Neuropsychiatry, 7*(2), 113–137.

Coltheart, M., Menzies, P. & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry, 15*, 261-287.

Fine, C., Gardner, M., Craigie, J. & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry,* *12*, 46-77.

Freeman, D. & Garety, P. A. (2014). Advances in understanding and treating persecutory delusions: a review. Social *Psychiatry* and *Psychiatric Epidemiology*, 49(8), 1179-1189.

Furl, N. & Averbeck, B. B. (2011). Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions. *The Journal of Neuroscience, 31*(48), 17572–17582.

Garety, P. A. & Freeman, D. (2013). The past and future of delusions research: From the inexplicable to the treatable. *The British Journal of Psychiatry,* *203*, 327-333.

Huq, S. F., Garety, P. A. & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *Quarterly Journal of Experimental Psychology A, 40*(4), 801-812.

Kahneman, D. (2011). Thinking, fast and slow: Farrar, Straus and Giroux.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237-51.

Langdon, R. (2011). The cognitive neuropsychiatry of delusional belief. *Wiley Interdisciplinary Reviews: Cognitive Science,* *2*, 449-60.

Langdon, R. & Coltheart, M. 2000: The cognitive neuropsychology of delusions. *Mind & Language, 15*(1), 183-216.

Lewis, D. K. (1973). *Counterfactuals.* Oxford: Basil Blackwell.

McKay, R. (2012). Delusional inference. *Mind & Language, 27*(3), 330-355.

Menon, M., Pomarol-Clotet, E., McKenna, P. & McCarthy, R. (2006). Probabilistic reasoning in schizophrenia: A comparison of the performance of deluded and non-deluded schizophrenic patients and exploration of possible cognitive underpinnings.

*Cognitive Neuropsychiatry, 11*, 521–536.

Mellet, E., Houde Ï, O., Brazo, P., Mazoyer, B., Tzourio-Mazoyer, N. & Dollfus, S. (2006). When a schizophrenic deficit becomes a reasoning advantage. *Schizophrenia Research, 84*, 359–364.

Parrott, M. & Koralus, P. (this issue). The erotetic theory of delusional thinking. *Cognitive Neuropsychiatry.*

Speechley, W. J., Whitman, J. C., & Woodward, T. S. (2010). The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. *Journal of Psychiatry and Neuroscience, 35*, 7-17.

Stone, T. & Young, A. W. (1997). Delusions and brain injury: The philosophy and psychology of belief. *Mind and Language, 12*, 327-364.

Van der Leer, L., Hartig, B., Goldmanis, M. & McKay, R. (2015). Delusion-proneness and ‘jumping to conclusions’: Relative and absolute effects*. Psychological Medicine, 45*(6), 1253-62.

1. “Cognitive neuropsychiatry… adopts a ‘levels-of-explanation’ approach to the study of psychiatric symptoms such as delusions and hallucinations… model[ing] the clinical phenomenology of *specific* symptoms in terms of disruption to normal processing of information about self and the world” (Langdon, 2011, p. 449, italics in original). [↑](#endnote-ref-1)
2. The distinction between externally stimulated and self-initiated questions is comparable to the well-known distinction between Type 1 (“intuitive”) processing and Type 2 (“analytic”) processing (e.g., Kahneman, 2011). A number of authors have suggested that the cognitive style of deluded individuals can be characterised as insufficiently analytic (e.g., Aimola Davies & Davies, 2009; Freeman & Garety, 2014). As a result, they tend not to override (“question”?) their automatic reactions to stimuli. [↑](#endnote-ref-2)
3. “Distinguish two kinds of parsimony… qualitative and quantitative. A doctrine is qualitatively parsimonious if it keeps down the number of fundamentalIy different *kinds* of entity: if it posits sets alone rather than sets and unreduced numbers, or particles alone rather than particles and fields, or bodies alone or spirits alone rather than both bodies and spirits. A doctrine is quantitatively parsimonious if it keeps down the number of instances of the kinds it posits; if it posits 1029 electrons rather than 1037, or spirits only for people rather than spirits for all animals. I subscribe to the general view that qualitative parsimony is good in a philosophical or empirical hypothesis; but I recognize no presumption whatever in favor of quantitative parsimony” (Lewis, 1973, p. 87). [↑](#endnote-ref-3)
4. In this respect their account seems inconsistent (they vacillate between endorsing a qualitative difference and endorsing a merely quantitative difference). An alternative way of reading P&K is as saying that *aside from their key hypothesis*, their account is qualitatively parsimonious (at one point they say “A notable virtue of this explanation is that it does not posit any cognitive states or operations that are radically different from that of normal subjects, *with the exception of the notion that delusional subjects are less inquisitive*” [my italics]). However, while on this interpretation P&K avoid inconsistency, they relinquish any real claim to qualitative parsimony. [↑](#endnote-ref-4)
5. “Damage to this normal safety-check mechanism (a deficit) is necessary, we think, to explain the presence of delusional beliefs” (Langdon & Coltheart, 2000, p. 203-4). [↑](#endnote-ref-5)
6. According to P&K, my account of Capgras delusion (McKay, 2012) “claims the subject’s prior in [the impostor] hypothesis is low and *therefore* discounted in her subsequent reasoning” (my italics). This is not quite accurate. While I do suggest that deluded individuals underweight their prior beliefs (and thus adopt beliefs that best explain the evidence available to them), this is not *because* those prior beliefs are low (or high). My point is that even very low prior beliefs (e.g., in the existence of impostors) can yield high posterior beliefs if the conditionalization process is distorted in the way I suggest. [↑](#endnote-ref-6)
7. Note that on the account of Adams et al., deluded individuals are over-responsive to sensory evidence, but do not necessarily depart from Bayesian reasoning. Instead, deluded individuals encode the precision of sensory evidence in an aberrant fashion (such that, relative to the precision of prior beliefs, the precision of sensory evidence is increased). The sensory evidence may, however, be combined with prior beliefs in a more-or-less Bayesian fashion. [↑](#endnote-ref-7)
8. One might object that there are plausible medical hypotheses that explain this evidence at least as well as the impostor hypothesis: for example, the hypothesis “This person is my loved one and I have suffered a stroke”. However, it is not obvious that this “stroke” hypothesis explains the evidence as well as (or better than) the impostor hypothesis. For one thing, the stroke hypothesis is quite general (P&K’s example “I am misperceiving due to illness” is even more general). A stroke can cause any number of psychological and physical problems. Given that one has had a stroke, the likelihood of any specific impairment (an impairment in the autonomic response to familiar faces, say) may be quite low. A more specific hypothesis would be “This person is my wife and I have had a stroke that has disconnected my face recognition system from my autonomic nervous system”. However, on the one hand the patient may not realize that an impairment in the autonomic response to familiar faces is even a possible consequence of stroke. So the patient might not even generate the more specific hypothesis. But on the other hand patients who *do* generate this hypothesis may not come to the attention of delusion researchers: “We assume that many people with similar brain injuries actually construct less fantastic accounts – they might say… that their vision is funny, etc. But these people do not get such a lot of attention from the medical profession!” (Stone and Young, 1997, p. 338). [↑](#endnote-ref-8)
9. P&K claim that the difference between control and delusion participants in the “classical” (unincentivized) version of the Beads Task “is explained by *particularly strong* Bayesian rationality” (their italics). The authors suggest that delusional participants who decide when the posterior probability of one of the jars is .97 are performing “roughly as a Bayesian algorithm would.” In my view there is no basis for this claim. The “Bayesian algorithm” simply provides the probabilities of events – in the absence of relevant costs, the algorithm cannot imply anything about when to decide. Moreover, the notion that .97 is some kind of all-purpose rational stopping point is easily contestable: with a gun to my head, I would be happy to sample beads all day rather than stop at .97 and risk a three per cent chance of losing my life. Obviously the stakes are far lower in the standard beads task, but the point is that without knowing what costs a participant anticipates, one cannot infer anything about the rationality or otherwise of their decision in this task. [↑](#endnote-ref-9)