**Measles, Magic and Misidentifications**

**A Defence of the Two-Factor Theory of Delusions**

**Ryan McKay**

Department of Psychology

Royal Holloway, University of London

Egham, Surrey TW20 0EX, United Kingdom

Email: [Ryan.McKay@rhul.ac.uk](mailto:Ryan.McKay@rhul.ac.uk)

Tel: +44 (0) 1784 414 993‬

**Abstract**

Corlett (this issue) provides a robust critique of the two-factor theory of delusions. The heart of his critique is two challenges he derives from a paper by Tranel and Damasio (1994), who illuminate the autonomic responses and brain damage of four patients often cited in support of the two-factor theory of Capgras delusion. I defend the two-factor theory against Corlett’s two key challenges, arguing that his first challenge has been previously addressed, and that his second challenge is overstated. In my view, these challenges do not negate the two-factor account. Nevertheless, two-factor theorists – and computational psychiatrists – should continue to devise and test falsifiable predictions of their respective theories.

**Keywords:** Delusions; Two-Factor Theory; Misidentification Delusions; Capgras Delusion.

In his article “Factor One, Familiarity and Frontal Cortex…” Phil Corlett (this issue) presents a bruising critique of the two-factor theory of delusions. In Corlett’s view the theory is “an erroneous summary of rather scant evidence”, “unmolested by direct empirical examination”; its “acolytes” not so much scientists as custodians of a kind of holy relic (to which they attribute “magical properties”).

I am not a neutral observer here, having co-authored some of the papers that Corlett critiques (Coltheart, Langdon & McKay, 2011; Ross, McKay, Coltheart & Langdon, 2016). On the other hand, I tend not to see the two-factor approach and Corlett’s own favoured theoretical approach – which I admire – as mutually exclusive (see McKay, 2012; McKay & Furl, 2017). I also think Corlett makes some valid points that will further the debate. So, while I aim to mount a robust defence of the two-factor theory against Corlett’s main arguments, I hope that, overall, my response will be taken as constructive rather than combative. But first, a quick recap of the logic of the two-factor theory.

***The Need for a Second Factor: An Example and an Analogy***

One may (a) interpret data falsely, but also (b) receive false data for interpretation.

*~ Elmer Ernest Southard, 1912, p. 328.*

In October 1879, a young woman was admitted to The State Lunatic Asylum at Danvers, Massachusetts. She complained of a noise in her head, which she attributed to the presence of bees in her skull. According to the pioneering neuropsychiatrist Elmer Ernest Southard, a pathologist at Danvers, this patient exemplified the *crystallization of* [at least some] *delusions around sensorial data of an abnormal sort* (Southard, 1912, p. 339, emphasis in original). For Southard, such patients often had bodily lesions that generated “false data for interpretation” (1912, p. 328); “the data of reasoning are as it were poisoned at the sensory source” (1916, p. 429).

The idea that delusions arise when individuals attempt to “interpret false data” has been incorporated in several subsequent theories of delusions. Maher, for example, viewed delusions as reasonable attempts to explain intense and anomalous experiences: “the locus of the pathology is in the neuropsychology of experience” (Maher, 1999). Two-factor theorists, however (e.g., Coltheart, 2010; Coltheart, Langdon & McKay, 2011; Davies, Coltheart, Langdon & Breen, 2001), view deluded individuals both as receiving false data for interpretation (factor one, which furnishes the content of the delusion) *and* as interpreting data falsely (factor two). Factor two is necessary, they claim, because some patients who appear to “receive false data for interpretation” are not delusional. The two-factor theory is thus based on the apparent dissociation between “false data” and “false belief”. The woman who believed bees were inside her skull, for instance, heard “a noise in her head” – but millions of people worldwide hear noises in their heads[[1]](#footnote-1) without coming to believe that their skulls are filled with live bees. So, something must explain why she – but not they – arrived at the delusion (Langdon & Connaughton, 2013). For two-factor theorists, that “something” is a deficient ability to evaluate candidate explanations of false data – the eponymous factor two.

By way of analogy, consider a two-factor model of viral infections. Factor one is exposure to a virus. The nature of the virus (e.g., measles, mumps) will furnish the “content” of the infection (the symptoms: perhaps a red, flat rash, or inflammation of the major salivary glands). But, of course, exposure to a virus may not be sufficient for an infection to develop, as one may have natural immunity or have been vaccinated against the disease. So, a “second factor” (lack of specific immunity, or generalised immunodeficiency) is needed to explain why some people exposed to the virus become infected while others do not.[[2]](#footnote-2)

***Corlett’s Two Challenges***

As Corlett acknowledges, two-factor accounts have been put forward for a range of monothematic delusions, including Frégoli delusion and mirrored-self misidentification. The focus of his critique, however, is the two-factor account of a different misidentification delusion: Capgras, the impostor delusion (which Corlett calls “the flagship of the two-factor explanatory fleet”).

For two-factor theorists, the “false data” in the Capgras case is a deficient autonomic response to familiar faces (indexed by skin conductance response [SCR]). The apparent dissociation between these “false data” and “false belief” comes from the fact that patients with damage to ventromedial prefrontal cortex (vmPFC) have also been shown to have deficient skin conductance responses to familiar faces ­– but these patients are not delusional (Tranel, Damasio & Damasio, 1995).

Corlett, however, has recently unearthed a paper by Tranel and Damasio (1994), “uncited by two-factor theorists”, which documents the full extent of brain damage of the four vmPFC patients often cited in support of the two-factor theory of Capgras delusion. The heart of Corlett’s critique is the two key conclusions he draws from the information in this paper.

Challenge to factor one: Corlett notes that the four patients lacked normal SCRs not just to familiar faces (as per Tranel et al., 1995) but to psychologically salient visual stimuli more generally (affectively laden pictures such as nudes or mutilated bodies). From this, he concludes that the four patients viewed by two-factor theorists as having the first factor in Capgras delusion, but not the second, “do not have a specific enough deficit to determine the Capgras delusion content”.

Challenge to factor two: Corlett notes that two of these four patients had damage to right dorsolateral prefrontal cortex (rDLPFC), from which he concludes that “rDLPFC damage is not sufficient for belief evaluation dysfunction as claimed by two-factor theorists”.

I will respond to each of these points in turn.[[3]](#footnote-3)

Response to Corlett’s first challenge:

Corlett’s challenge to factor one has actually been acknowledged and discussed by two-factor theorists on a number of occasions (Breen, Caine & Coltheart, 2000, 2002; Langdon & Coltheart, 2000). For example, in one of the earliest papers on the two-factor theory, Langdon and Coltheart (2000, p. 191) wrote the following:

But perhaps assuming that a similar pattern of empirical data indexes a similar perceptual anomaly in Tranel et al.’s (1995) frontal patients and Ellis et al.’s (1997a) Capgras patients is unjustified… It is worth noting, in this context, that the Tranel et al. (1995) study and the Ellis et al. (1997a) study differ in that the former did not demonstrate that the under-responsiveness of their patients was circumscribed. Indeed, Damasio, Tranel and Damasio (1991) have reported elsewhere that patients with bilateral ventromedial frontal lesions fail to show normal autonomic responses to emotionally charged visual stimuli such as pictures of mutilations and social disasters. It is therefore possible, that, even though Tranel et al.’s (1995) frontal patients show empirical evidence of a discordant mismatch between intact explicit face recognition and loss of autonomic face recognition, if that mismatch occurs in the context of general affective flatness (which may well be the case in these patients), then the resultant perceptual experience (when encountering a familiar face) may differ from that of Capgras patients and may indeed not even register as aberrant.

Interestingly, because of these doubts about whether the Tranel et al. (1995) patients represent appropriate first-factor analogues of Capgras patients, Langdon and Coltheart (2000) went on to discuss other cases that they viewed as potentially more compelling in this regard. This seems to me to undercut an implicit premise of Corlett’s overall argument – that whether the two-factor theory of delusions stands or falls depends on its success in explaining Capgras delusion specifically.

But if the vmPFC patients do not fit the bill, are there any non-delusional patients with a specific enough deficit to determine the content of Capgras delusion? Although it lacks an objective measure (e.g., SCR) to confirm common phenomenology, a case discussed by Turner and Coltheart (2010) is certainly suggestive in this regard. The patient had undergone neurosurgery to treat intractable epilepsy, and subsequently reported that her mother *felt* different: “[T]he first thing I noticed was Mum, when she walked in the room... it was like a picture of her, but it wasn't her... Just didn’t feel like her” (pp. 371-2). Nevertheless, this patient did not adopt the delusional belief that her mother had been replaced by someone physically identical.

Response to Corlett’s second challenge:

What about Corlett’s challenge to factor two? While the two-factor theory does not specifically *predict* the existence of non-delusional patients who have both the putative first factor in Capgras delusion (impaired autonomic responses to faces) and damage to rDLPFC, I don’t see that evidence of rDLPFC damage in such patients is a big problem for the theory. For one thing, rDLPFC is a huge and heterogeneous region of the brain, so even substantial damage to rDLPFC might not disrupt the circuits critical for belief evaluation, assuming rDLPFC is indeed the locus of those circuits. But in any case, claims about rDLPFC damage as the neuroanatomical substrate of factor two are relatively tentative (e.g., Coltheart et al., 2018[[4]](#footnote-4)). Most two-factor theorists have associated factor two more generally with dysfunction in the right hemisphere (e.g., Davies & Coltheart, 2000), the frontal lobe (e.g., Turner, Shores, Breen and Coltheart, 2017), right frontal cortex (e.g., Coltheart, 2007; Coltheart, Langdon & McKay, 2007) or right lateral prefrontal cortex (rLPFC[[5]](#footnote-5); e.g., Coltheart, 2010; Coltheart et al., 2011; Coltheart et al., 2018).

Later, Corlett claims that in order “to be more confident of two dissociable, independent, factors in delusions,” two-factor theorists need to produce “a patient with Capgras delusion who doesn’t have an SCR disruption”. But this is tantamount to demanding a patient with measles who hasn’t been exposed to the measles virus. What seems more reasonable is to ask for evidence of factor two in the absence of factor one. In other words, what independent evidence is there for the existence of factor two?

Without such evidence, one possibility – which Corlett articulates compellingly – is that the dissociation between anomalous experience and delusion that two-factor theory takes as its point of departure is more apparent than real. Putatively, the second factor is what bridges the gap between the vmPFC and Capgras cases; between tinnitus and the bees-in-the-head delusion; between déjà vu and its delusional manifestation, déjà vecu[[6]](#footnote-6); and so on. The assumption in such cases is that the anomalous experience is equivalent for the deluded and non-deluded individuals, e.g., the vmPFC and Capgras patients have received equivalent “false data for interpretation.” But as Corlett points out, “Damage to the perception/belief machinery… could simply be less extensive in the vmPFC cases.” In other words, it could be that the distinction between the vmPFC and Capgras cases is not that the vmPFC patients lack factor two, but that factor one (the *only* factor) is simply less pronounced in these patients; they just got a “smaller dose of the virus”. In Sakakibara’s (2018) recent analysis, where the patients may differ is in the *intensity* of their anomalous experience.[[7]](#footnote-7)

For my money, Darby, Laganiere, Pascual-Leone, Prasad and Fox’s (2017) recent lesion network mapping work (to which Corlett alludes) provides compelling evidence that two independent factors underpin misidentification delusions such as Capgras delusion. These authors used a resting-state connectivity analysis to identify two key regions functionally connected to the lesion sites in a sample of patients with misidentification delusions. In separate meta-analyses of previous fMRI studies, they found that the first of these regions (left retrosplenial cortex) was the region most activated by personally familiar (versus unfamiliar) stimuli, and that the second region (right frontal cortex) was activated by violations (versus confirmations) of participant expectations. These results thus indicate that misidentification delusions imply disruption to not one but two functionally distinct networks – a familiarity network (putatively factor one) and a belief evaluation network (putatively factor two; see McKay & Furl, 2017). Moreover, these findings suggest factor two can occur in the absence of factor one for Capgras, as the pattern of connectivity characterizing persecutory delusions implicated disruption to the belief evaluation network but not to the familiarity network.

***Conclusion***

None of this is to imply that two-factor theory can rest on its laurels. Two-factor theorists – and computational psychiatrists – should continue to devise and test falsifiable predictions of their respective theories. Corlett suggests, for example, that speculations about the nature of the first factor in delusions other than Capgras (e.g., Frégoli[[8]](#footnote-8)) should be put to the test, and I agree.[[9]](#footnote-9) But equally, Corlett should clarify how his preferred theoretical scheme accounts for the recurrence of certain delusional themes. And while two-factor theorists await opportunities to test patients with relevant monothematic delusions (some of which, of course, are rare), it seems to me that other empirical approaches – lesion network mapping (Darby & Fox, 2017; Darby et al., 2017), for instance, or work using hypnotic analogues of clinical delusions (e.g., Barnier et al., 2008; Coltheart et al., 2018; Connors, Barnier, Coltheart, Cox & Langdon, 2012; Connors, Cox, Barnier, Langdon & Coltheart, 2012; Cox & Barnier, 2010; Oakley & Halligan, 2013) – can provide invaluable evidence for refining our theories and potentially adjudicating between them.

**References**

Barnier, A. J., *et al.* (2008). Developing hypnotic analogues of clinical delusions: Mirrored-self misidentification. *Cognitive Neuropsychiatry, 13*(5), 406-30.

Brighetti, G., Bonifacci, P., Borlimi, R. & Ottaviani, C. (2007). ‘Far from the heart far from the eye’: Evidence from the Capgras delusion. *Cognitive Neuropsychiatry, 12*(3), 189–97.

Breen, N., Caine, D. & Coltheart, M. (2000). Models of face recognition and delusional misidentification: A critical review. *Cognitive Neuropsychology, 17*(1–3), 55–71.

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. (2007). The 33rd Sir Frederick Bartlett Lecture: Cognitive neuropsychiatry and delusional belief. *The Quarterly Journal of Experimental Psychology, 60*(8), 1041-62.

Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences, 1191*, 16-26. doi:10.1111/j.1749-6632.2010.05496.x

Coltheart, M., *et al.* (2018). Belief, delusion, hypnosis, and the right dorsolateral prefrontal cortex: A transcranial magnetic stimulation study. *Cortex, 101,* 234-48. doi: 10.1016/j.cortex.2018.01.001

Coltheart, M., Langdon, R. & McKay, R. (2007). Schizophrenia and monothematic delusions. *Schizophrenia Bulletin*, *33*(3), 642-647.

Coltheart, M., Langdon, R. & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, *62*, 271–98.

Connors, M. H., Barnier, A. J., Coltheart, M., Cox, R. E., & Langdon, R. (2012). Mirrored-self misidentification in the hypnosis laboratory: Recreating the delusion from its component factors. *Cognitive Neuropsychiatry, 17*(2), 151-76.

Connors, M. H., Cox, R. E., Barnier, A. J., Langdon, R. & Coltheart, M. (2012). Mirror agnosia and the mirrored-self misidentification delusion: A hypnotic analogue. *Cognitive Neuropsychiatry, 17*(3), 197-226.

Cook, J., Lewandowsky, S. & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. PLoS ONE, 12(5): e0175799. doi:10.1371/journal.pone.0175799

Corlett, P. (this issue). Factor One, familiarity and frontal cortex: A challenge to the two-factor theory of delusions. *Cognitive Neuropsychiatry.*

Corlett, P. R., *et al.* (2007). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain, 130*(9), 2387-2400.

Cox, R. E. & Barnier, A. J. (2010). Hypnotic illusions and clinical delusions: Hypnosis as a research method. *Cognitive Neuropsychiatry, 15*(1), 202-32.

Darby, R. R. & Fox, M. D. (2017) Reply: Capgras syndrome: Neuroanatomical assessment of brain MRI findings in an adolescent patient. *Brain, 140*, e44.

Darby, R. R., Laganiere, S., Pascual-Leone, A., Prasad, S. & Fox, M. (2017). Finding the imposter: Lesions causing delusional misidentifications are characterized by a unique pattern of brain connectivity. *Brain, 140*,497–507.

Davies, M. & Coltheart, M. (2000). Introduction: Pathologies of belief. *Mind & Language*, *15*(1), 1–46.

Davies, M., Coltheart, M., Langdon, R. & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry & Psychology, 8*(2-3), 133-58. doi:10.1353/ppp.2001.0007

Kalra, S., Chancellor, A. & Zeman, A. (2007). Recurring déjà vu associated with 5-hydroxytryptophan. *Acta Neuropsychiatrica, 19*, 311-13.

Kochkin, S., Tyler, R. & Born, J. (2011). MarkeTrak VIII: The prevalence of tinnitus in the United States and the self-reported efficacy of various treatments. *Hearing Review, 18*(12), 10-26.

Kucharski, A. (2016). Study epidemiology of fake news. *Nature, 540*(7634). 525.

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

Langdon, R. & Connaughton, E. (2013). The neuropsychology of belief formation. In F. Kreuger & J. Grafman (Eds.), *The Neural Basis of Human Belief Systems* (pp. 19-42). Hove; New York: Taylor & Francis.

Langdon, R., Connaughton, E. & Coltheart, M. (2014). The Fregoli delusion: A disorder of person identification and tracking. Topics in Cognitive Science, 6, 615–631.

Maher, B. A. (1988). Anomalous experience and delusional thinking: The logic of explanations. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs* (pp. 15-33). New York: Wiley.

Maher, B. A. (1999). Anomalous experience in everyday life: Its significance for psychopathology. *The Monist, 82*, 547-570.

McGuire, W. J. (1964). Some contemporary approaches. *Advances in Experimental Social Psychology, 1*, 191–229. doi:10.1016/ S0065-2601(08)60052-0.

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

McKay, R. & Furl, N. (2017). Doppelgängers and dissociations: Lesion network mapping illuminates misidentification delusions. *Brain, 140*(2), 262-265. doi:10.1093/brain/aww323.

Negro, E., *et al.* (2015) Neurofunctional signature of hyperfamiliarity for unknown faces. *PLoS ONE, 10*(7), e0129970.

Oakley, D. A. & Halligan, P. W. (2013). Hypnotic suggestion: Opportunities for cognitive neuroscience. *Nature Reviews Neuroscience*, *14*, 565–576.

Roozenbeek, J. & van der Linden, S. (2018). The fake news game: Actively inoculating against the risk of misinformation. Journal of Risk Research. doi:10.1080/13669877.2018.1443491

Ross, R. M., McKay, R., Coltheart, M. & Langdon, R. (2016). Perception, cognition, and delusion. *Behavioral and Brain Sciences, 39*, e258. doi: 10.1017/S0140525X15002691.

Sakakibara, E. (2018). Intensity of experience: Maher’s theory of schizophrenic delusion revisited. *Neuroethics*. doi:10.1007/s12152-018-9385-4

Southard, E. E. (1912). On the somatic sources of somatic delusions. *The Journal of Abnormal Psychology, 7*(5), 326-339.

Southard, E. E. (1916). On the application of grammatical categories to the analysis of delusions. *The Philosophical Review, 25*(3), 424-455.

Tranel, D. & Damasio, H. (1994). Neuroanatomical correlates of electrodermal skin conductance responses. *Psychophysiology, 31*(5), 427-438.

Tranel, D., Damasio, H. & Damasio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience, 7*, 425-432.

Turner, M. & Coltheart, M. (2010). Confabulation and delusion: A common monitoring framework. *Cognitive Neuropsychiatry, 15*, 346-376.

Turner, M. S., Shores, E. A., Breen, N. & Coltheart, M. (2017). Déjà vecu for news events but not personal events: A dissociation between autobiographical and non-autobiographical episodic memory processing. *Cortex, 87*, 142-155. doi: 10.1016/j.cortex.2016.11.006.

Vuilleumier, P., Mohr, C., Valenza, N., Wetzel, C. & Landis, T. (2003). Hyperfamiliarity for unknown faces after left lateral temporo‐occipital venous infarction: A double dissociation with prosopagnosia. Brain, 126(4), 889–907.

1. The authors of one prevalence study estimate that about 30 million Americans suffer from tinnitus (Kochkin, Tyler & Born, 2011). [↑](#footnote-ref-1)
2. Many others have invoked an analogy between resistance to false data (e.g., “fake news”) and biological resistance (e.g., Cook, Lewandowsky & Ecker, 2017; Kucharski, 2016; McGuire, 1964; Roozenbeek & van der Linden, 2018). [↑](#footnote-ref-2)
3. As an aside, it’s worth noting that these are not really two independent, mutually reinforcing challenges. On the contrary, Corlett’s first challenge – if successful – somewhat undermines his second. After all, if the Tranel et al. (1995) patients don’t actually have the first factor for Capgras delusion (as Corlett alleges), then it hardly matters whether they have the second factor or not – either way, we would not expect them to have Capgras delusion. [↑](#footnote-ref-3)
4. “[T]he evidence seems clear that damage to right frontal cortex is associated with the presence of delusional belief, and even that the specific region of right frontal cortex that is critical here is rLPFC ­– and *possibly* an even more specific region, rDLPFC” (Coltheart et al., 2018, p. 237, my emphasis). [↑](#footnote-ref-4)
5. Ironically, this more specific prediction is based, in part, on Corlett’s own work (Corlett et al., 2007). [↑](#footnote-ref-5)
6. Turner et al. (2017) contrast a case of medication-induced non-delusional déjà vu (“I was a little freaked out when I watched TV as I felt I was watching repeats, although I knew I wasn't”; Kalra, Chancellor & Zeman, 2007, p. 312) with their own case of déjà vecu (“Everyone says ‘you only think you've seen it before’. But I'll swear black and blue that I have seen it before.”; Turner et al., 2017, p. 144). [↑](#footnote-ref-6)
7. To return to the bees-in-the-head case, it’s worth noting that whereas the most common cause of tinnitus is cochlear damage, the Danvers patient with the bees delusion was thought to have a condition entailing direct mechanical stimulation of auditory cortex (see Maher, 1988). Her resulting auditory experience may well have been more intense than the sounds experienced by the average sufferer of tinnitus. [↑](#footnote-ref-7)
8. Individuals with Frégoli delusion believe that strangers in their environment are actually familiar people in disguise (see Langdon, Connaughton & Coltheart, 2014). Corlett notes that “the control condition (the vmPFC case equivalent) for Fregoli [has not] been identified”. There are, however, non-delusional individuals with intact face recognition who report an abnormal feeling of familiarity for unknown faces (e.g., Vuilleumier, Mohr, Valenza, Wetzel & Landis, 2003; Negro et al., 2015). [↑](#footnote-ref-8)
9. Corlett also suggests confirming that Capgras patients have deficient autonomic responses to *personally* familiar faces (“a dearth of responding to famous people is not enough”), but there is already evidence they do (Brighetti, Bonifacci, Borlimi & Ottaviani, 2007). [↑](#footnote-ref-9)