

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/123679>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

This is the Accepted Manuscript. The Version of Record of this manuscript has been published and is available in *Cognitive Neuropsychiatry* on 04 Aug 2019  
<https://www.tandfonline.com/doi/full/10.1080/13546805.2019.1652156>

# **Continuing commentary: Challenges or misunderstandings? A defence of the two-factor theory against the challenges to its logic**

Chenwei Nie

*Department of Philosophy, University of Warwick, Coventry, UK*

Email: [c.nie.1@warwick.ac.uk](mailto:c.nie.1@warwick.ac.uk)

## **ABSTRACT**

Corlett (2019) raises two groups of challenges against the two-factor theory of delusions: One focuses on weighing “the evidence for ... the two-factor theory”; the other aims to question “the logic of the two-factor theory” (ibid., p. 166). McKay (2019) has robustly defended the two-factor theory against the first group. But the second group, which Corlett believes is in many aspects independent of the first group and Darby (2019, p. 180) takes as “[t]he most important challenge to the two-factor theory raised by Dr. Corlett”, has by large remained. Here I offer my two cents in response to the second group. More specifically, I argue that Corlett’s challenges to the logic of the two-factor theory, concerning modularity, double dissociation and cognitive penetration, seem to be based on some misunderstandings of the two-factor theory.

**KEYWORDS:** delusions; two-factor theory; modularity; double dissociation; cognitive penetration

It is time to overturn the received wisdom.

~ Martin Davies, 2010, p. 502

Corlett’s (2019) intriguing discussion covers many important issues, including modularity, double dissociation and cognitive penetration. He argues that the logic of the two-factor theory of delusions related to these issues is problematic. In the following, I shall begin with

Corlett's conception of the two-factor theory, then discuss Corlett's challenges. After each challenge, I shall give my response in defence of the two-factor theory.

### **Corlett's conception of the two-factor theory**

In essence, the two-factor theory argues that two factors are needed to explain delusions. The first factor is an anomaly in the perceptual system and the second factor is a deficit in the belief evaluation system (e.g., Coltheart, Langdon, & McKay, 2011; Davies & Coltheart, 2000; Davies, Coltheart, Langdon, & Breen, 2001; Davies & Egan, 2013; Davies, McGill, & Davies (in press); Langdon & Coltheart, 2000; Nie, 2016).<sup>1</sup> According to Corlett, the two-factor theory makes the following assumptions as well:

- (1) The first factor and the second factor are independent modules.<sup>2</sup>
- (2) The independence between perception and belief evaluation requires the support of double dissociation evidence.
- (3) Cognitive penetration is not permissible.

Overall, they constitute what Corlett (2019, p. 166) calls the "logic of the two-factor theory".

---

<sup>1</sup> The two-factor theory is a family of accounts. They share this commitment but their views on the details of the two factors may vary. Though it seems that Corlett aims to challenge the two-factor theory in general, only a selection of papers mainly (co-)authored by Coltheart is mentioned, which represents a particular two-factor account. To make this commentary more straightforward, I shall focus on the papers cited by Corlett. But, I think, the arguments and explanations apply to the two-factor theory in general. (Thanks to an anonymous reviewer for helping me clarify this point.)

<sup>2</sup> Corlett (2019, p. 171) uses expressions like "factor-2, ... is a module". This might be misconstrued as meaning that the deficit in the belief evaluation system is a module. I think what Corlett actually means by this expression is that the impaired belief evaluation system is a module. Nevertheless, for the sake of argument, this commentary shall follow Corlett's use.

He directly challenges these three assumptions.<sup>3</sup>

### **Corlett's first challenge to Assumption (1): a puzzle**

The problem with Assumption (1), according to Corlett, is that there is a puzzle in assuming that the second factor is a module.<sup>4</sup> The puzzle is that, on the one hand, Coltheart (1999) argues that modules are domain-specific, on the other hand, Corlett (2019, p. 171) thinks, "Coltheart insists" that the second factor, which is not domain-specific, is a module.

### **Response**

It is correct that by modules Coltheart means domain-specific systems. However, the claim that Coltheart insists on the modularity of the second factor seems to be a misunderstanding. When Coltheart (2002) talks about delusions and the belief evaluation system, he emphasizes that the belief evaluation system is not modular. Following Fodor (1983), Coltheart's concept of modularity is about low-level cognitive abilities such as visual perception, rather than high-level cognitive abilities such as belief evaluation. Besides, Coltheart makes it clear that the second factor corresponds to Fodor's non-modular central system when he relates the

---

<sup>3</sup> Darby (2019, p. 180) gives a different presentation of the assumptions Corlett argues against: "(1) Modularity of brain functions, (2) Unidirectional flow of information, and (3) Cognitive processes that cannot influence perceptual processes." He does not include, what I call, Assumption (2) because, I think, his intention is not to give a summary of Corlett's arguments but to provide "additional data" to support the related arguments (ibid., p. 178). I choose not to elaborate on, what I call, Assumption (3) because, as I shall argue later, the related discussions in both Corlett and Darby's articles are mainly general statements which seem too ambiguous to be a clear argument against the two-factor theory.

<sup>4</sup> Corlett repeatedly claims that the modularity of mind and brain is a problem for the two-factor theory. But it is unclear whether he is also denying that in monothematic delusions the first factor is modular in Coltheart's sense, that is the first factor is a domain-specific deficit (for a discussion on the empirical evidence, see also McKay, 2019).

difficulty of understanding the nature of the second factor to the difficulty of understanding the nature of Fodor's central system:

This kind of attempt to explain delusions in cognitive-neuropsychological terms is particularly challenging precisely because cognitive psychology does not currently offer an adequate theory of the normal processes of belief formation and evaluation; perhaps Fodor's first law is correct, which means that such a theory will never be found, in which case cognitive neuropsychiatry will never flourish. (Coltheart, 2002, p. 171)

The two-factor theory proposes that the second factor is different from the first factor and may also accept that the first factor is a module in monothematic delusions (Coltheart, 2013). This, I suspect, is misconstrued as suggesting the second factor is a different module. Once the second factor is properly understood as a deficit in the non-modular belief evaluation system, the puzzle is cracked.

### **Corlett's second challenge to Assumption (1): an outdated view**

Another issue with modularity is that, both Corlett and Darby seem to think, the two-factor theory is committed to an outdated view according to which a mental function is not only implemented in a certain neuroanatomical module, but more importantly, a neuroanatomical module is restricted to a limited and specific *region*. They vigorously refute this view, taking it as a refutation of the two-factor theory as well. Corlett argues that “[m]odern network neuroscience [rejects] ... rigidly ascribing functions to particular *regions*” (Corlett, 2019, p. 173, emphasis added). Similarly, Darby argues that “numerous studies now show that complex brain processes result from the function of integrated brain networks rather than discrete brain *regions*” (Darby, 2019, p. 180, emphasis added).

### **Response**

In this challenge, Corlett and Darby seem to be concerned with the two-factor theory's

proposal that the second factor is “neurally realized in right dorsolateral prefrontal cortex (rDLPFC)” which is a particular brain region (Coltheart et al., 2018, p. 234). The two-factor theory indeed argues that this proposal is supported by current evidence about *the second factor in delusions*. But it would be a misunderstanding to say that the two-factor theory has any a priori theoretical commitment to the view that *all mental functions* should be ascribed to particular regions. In fact, in one of the papers cited by Corlett, Coltheart and colleagues write explicitly that the belief evaluation system can involve different brain regions:<sup>5</sup>

Current evidence suggests that [the impairment of belief evaluation] will be associated with right lateral prefrontal cortex pathology, but the belief evaluation system itself must involve a number of different cognitive subcomponents, and damage to any one of these subcomponents might disrupt inputs into the critical functioning of the right lateral prefrontal region. These different subcomponents will, of course, be realized in the brain as a neural circuit involving *various brain regions*, only one of which might be located in right lateral prefrontal cortex. (Coltheart et al., 2011, p. 286, emphasis added)

To challenge the rDLPFC proposal, merely talking about the apparent similarity between the proposal and the “outdated” view in ascribing a mental function to a particular region is not enough. Rather, one needs to engage with the overall evidence provided by the two-factor theory (Coltheart 2007; Coltheart et al., 2018). This is however absent in Corlett and Darby’s challenge (for a defence against Corlett’s challenge based on one particular piece of counter-evidence, see McKay, 2019).

### **Corlett’s challenges to Assumption (2): a lack of evidence and a paradox**

The problems with Assumption (2), according to Corlett, are: first there is no reported double dissociation evidence; second the assumption of double dissociation evidence would create a

---

<sup>5</sup> Nie (2016) further develops the idea that the belief evaluation system involves not only different brain regions but also different functions.

“paradox”. The double dissociation evidence in Corlett’s conception seems to be as follows:

- (*Capgras*) patient A has a deficit in perception but no deficit in belief evaluation.<sup>6</sup>
- *Capgras* patient B has a deficit in belief evaluation but no deficit in perception.

This is why Corlett thinks that “[i]f Coltheart wishes ... to be more confident of two dissociable, independent, factors in delusions, ...he needs to find a patient with *Capgras* delusions who doesn’t have” a deficit in perception (Corlett, 2019, p. 169).<sup>7</sup> Call this kind of evidence Corlett’s Double Dissociation Evidence (CDDE). The first problem, Corlett then argues, is that *Capgras* patient B which is necessary for CDDE has not been reported. The second problem, Corlett further argues, is that even if there were a reported *Capgras* patient B (and the two-factor theory might hence enjoy the fulfilment of the double dissociation evidence), the *Capgras* patient B who has *only one* deficit in belief evaluation would “paradoxically ... obviate the two-factor theory” which requires *two* independent deficits (ibid.).

## Response

In the first place, it is unclear why Corlett thinks that the lack of double dissociation between perception and belief evaluation is a problem for the two-factor theory. Here it would be helpful to make a distinction between the central claim of the two-factor theory, that is there are two *conceptually* distinct factors in delusions, and, what Davies and Egan (2013, p. 691)

---

<sup>6</sup> It is not clear whether or not subject A in Corlett’s conception needs to be a *Capgras* patient. In any case, the discussion focuses on subject B.

<sup>7</sup> McKay seems to find Corlett’s demand of *Capgras* patient B confusing. He writes that “this is tantamount to demanding a patient with measles who hasn’t been exposed to the measles virus” (McKay, 2019, p. 186). Here, I’m trying to make plain Corlett’s consideration behind the demand.

call, the dissociation argument for the two-factor theory, which says that a second factor is needed to explain delusions because there are cases in which the subjects have the first factor but not the delusion. Corlett's paper begins with an introduction of the dissociation argument. Does Corlett mean that the lack of double dissociation evidence is a problem for the dissociation argument? This cannot work because the dissociation argument depends on single, rather than double, dissociation evidence (for a review of the single dissociation evidence in monothematic delusions, see Coltheart, Langdon & McKay, 2011).

Is the lack of double dissociation evidence a problem for the central claim? Admittedly, some kind of double dissociation evidence can support the central claim. But it is not necessary for the claim because, first of all, the current two-factor claim is based on single, rather than double, dissociation evidence. More importantly, the two-factor claim itself is compatible with various possible relationships between the first factor and the second factor at either the *functional* level or the *neuroanatomical* level. Consider a possibility according to which it is an empirical fact, probably due to the neuroanatomical overlap, that the first factor and the second factor always appear together in delusions. Consider another possibility according to which the first factor of delusions is so anomalous that it will impede the execution of our rational ability (for a similar idea, see Hoerl, 2018). Consider yet another possibility according to which the second factor contributes to the formation of the first factor or they have some degree of mutual influence. In these possibilities, even though there is no double dissociation between the first factor and the second factor, the central claim obtains.<sup>8</sup>

---

<sup>8</sup> I think even the single dissociation evidence is not necessary for the two-factor claim. Admittedly, the *current* two-factor theory is mainly based on the dissociation argument. In these possibilities, the *possible* two-factor theories would be based on some different arguments from the dissociation argument. Nonetheless, in both cases the central claim obtains. (I'm grateful to an anonymous

In short, the lack of double dissociation evidence is neither a problem for the dissociation argument nor a problem for the central claim. In other words, it is not a problem for the current two-factor theory and would not be a problem for the possible two-factor theories in these possibilities.

Now let us consider what kind of double dissociation evidence is needed if there were someone who wanted a double dissociation between perception and belief evaluation. Let us also put aside the issue that double dissociation is typically about two low-level cognitive abilities and does not involve high-level belief evaluation. In my understanding, it should be:

- Subject A has a deficit in perception but no deficit in belief evaluation.
- Subject B has a deficit in belief evaluation but no deficit in perception.

Call this kind of evidence the Simple Double Dissociation Evidence (SDDE). The main difference between SDDE and CDDE is that subject B in the latter is a Capgras patient. Corlett's claim is correct that CDDE would be a problem for the two-factor theory because it contains a deluded subject with only one deficit rather than two. But, SDDE is compatible with the two-factor theory because it does not contain a deluded subject with only one deficit. That is to say, even if a two-factor theorist wanted some double dissociation evidence, she could simply look for the unproblematic SDDE. Corlett provides no reason why the two-factor theorist has to ask for the paradoxical CDDE.<sup>9</sup>

### **Corlett's challenge to Assumption (3): another outdated view**

Corlett (2019) and Darby (2019) take the two-factor theory as rejecting top-down influences

---

reviewer for raising the concern about the relationship between the dissociation argument and the two-factor theory.)

<sup>9</sup> Thanks to an anonymous reviewer for pressing me to clarify the point in this paragraph.

from belief to perception. This is a problem because, they argue, many recent works have shown that there is some degree of top-down influence.

## **Response**

Corlett admits that the rejection of top-down influence is not an explicit claim of the two-factor theorists. But why does Corlett still think that it is implicit in the two-factor theory? He writes:

[T]wo-factor theory is not *explicit* about cognition penetrating perception, but two factor theorists claim that perception and belief are dissociated modules, with a uni-directional flow of information from perception to belief (else only one factor damage to top-down belief effects—would suffice ...), so the rejection of cognitive penetration is *implicit* and I believe essential to two-factor theory as espoused by Coltheart and colleagues. (Corlett, 2019, p. 173, emphasis added)

It looks to me that Corlett might have two reasons. The first reason seems to be that the two-factor theory assumes that perception and belief evaluation are double-dissociable modules. I have argued that this is not true. The second reason seems to be that if the two-factor theory allows top-down influences, then “one only factor damage”, rather than two factors, “to top-down belief effects ... would suffice”. It is not clear what Corlett means by “one only factor damage” and why it is a problem for the two-factor theory. If it means neuroanatomical damage, such as the one proposed by Darby, Laganier, Pascual-Leone, Prasad, and Fox (2017), then it is consistent with the two-factor theory in the sense that the two-factor theory allows that the perceptual deficit and the belief evaluation deficit at the functional level share same neuroanatomical damage (McKay & Furl, 2017). If it means functional damage, say, some damage to the prediction-error mechanism, then it is also not straightforwardly against the two-factor theory (for a view arguing that the prediction-error theory and the two-factor theory are compatible, see Miyazono, Bortolotti, & Broome, 2015). In short, the claim that

the two-factor theory implicitly rejects cognitive penetration seems to be unfounded.

## **Conclusion**

In his classic paper on double dissociation, Davies reminds us to be cautious of the received wisdom about a theory which is some inaccurate assumption that “many people believe—or many people believe that many people believe” the theory is based on (Davies, 2010, p. 501). If my analysis so far is along the right lines, then it seems that Corlett’s challenges to the logic of the two-factor theory are mainly about the questionable assumptions which are misconstrued as the basis of the two-factor theory. Once the received wisdom about the two-factor theory is identified and overturned, the challenges should be gone with it. Despite disagreeing with Corlett’s challenges to the two-factor theory, I believe that Corlett’s own prediction-error theory looks to be a viable account with many distinguished contributions. Taking different theories together, I think, would be a better way to further advance our understanding of delusions.

## **Acknowledgements**

This work is supported by a Chancellor's International Scholarship from the University of Warwick. I thank Johannes Roessler for his invaluable suggestions, thank Ryan McKay for a recent discussion about the two-factor theory, thank Ryan Darby for sharing Darby et al.’s (2017) supplementary materials by personal communication, and thank Anthony David and the anonymous reviewers for their very helpful comments. I’m responsible for the ideas in it.

## **Disclosure statement**

No potential conflict of interest was reported by the author.

## **References**

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, 3(3), 115-120.  
doi: 10.1016/S1364-6613(99)01289-9

- Coltheart, M. (2002). Cognitive Neuropsychology. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology (Vol. 4: Methodology in experimental psychology)* (3rd ed., pp. 139-174). Hoboken, NJ: Wiley.
- Coltheart, M. (2007). The 33rd Bartlett lecture: Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology*, 60(8), 1041-1062. doi: 10.1080/17470210701338071
- Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, 62, 271-298. doi: 10.1146/annurev.psych.121208.131622
- Coltheart, M. (2013). On the distinction between monothematic and polythematic delusions. *Mind & Language*, 28(1), 103-112. doi: 10.1111/mila.12011
- Coltheart, M., Cox, R., Sowman, P., Morgan, H., Barnier, A., Langdon, R., ... Polito, V. (2018). Belief, delusion, hypnosis, and the right dorsolateral prefrontal cortex: a transcranial magnetic stimulation study. *Cortex*, 101, 234-248. doi: 10.1016/j.cortex.2018.01.001
- Corlett, P. R. (2019). Factor one, familiarity and frontal cortex: A challenge to the two-factor theory of delusions. *Cognitive Neuropsychiatry*, 24(3), 165-177. doi: 10.1080/13546805.2019.1606706
- Darby, R. R. (2019). A network-based response to the two-factor theory of delusion formation. *Cognitive Neuropsychiatry*, 24(3), 178-182. doi: 10.1080/13546805.2019.1606709
- Darby, R. R., Laganier, S., Pascual-Leone, A., Prasad, S., & Fox, M. D. (2017). Finding the imposter: Brain connectivity of lesions causing delusional misidentifications. *Brain*, 140(2), 497-507. doi: 10.1093/brain/aww288
- Davies, M. (2010). Double dissociation: Understanding its role in cognitive neuropsychology. *Mind & Language*, 25(5), 500-540. doi: 10.1111/j.1468-0017.2010.01399.x
- Davies, M., & Coltheart, M. (2000). Introduction: Pathologies of belief. *Mind & Language*, 15(1), 1-46. doi: 10.1111/1468-0017.00122
- Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, & Psychology*, 8(2), 133-158. doi: 10.1353/ppp.2001.0007
- Davies, M., & Egan, A. (2013). Delusion, cognitive approaches: Bayesian inference and compartmentalisation. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J.

- Sadler, G. Stanghellini & T. Thornton (Eds.), *The oxford handbook of philosophy and psychiatry* (pp. 689-727). Oxford: Oxford University Press.
- Davies, M., McGill, C. L., & Aimola Davies, A. M. (in press). Anosognosia for motor impairments as a delusion: Anomalies of experience and belief evaluation. In A. L. Mishara, P. R. Corlett, P. C. Fletcher, A. Kranjec, & M. A. Schwartz (Eds.), *Phenomenological neuropsychiatry: How patient experience bridges clinic with clinical neuroscience*. New York: Springer.
- Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, Mass.: MIT Press.
- Hoerl, C. (2018). Understanding and explaining in psychiatry. In S. Giovanni, B. Matthew, R. Andrea, F. Anthony Vincent, F.-P. Paolo & R. René (Eds.), *The oxford handbook of phenomenological psychopathology*. Oxford: Oxford University Press.
- Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language*, 15(1), 184-218. doi: 10.1111/1468-0017.00129
- McKay, R. (2019). Measles, magic and misidentifications: A defence of the two-factor theory of delusions. *Cognitive Neuropsychiatry*, 24(3), 183-190. doi: 10.1080/13546805.2019.1607273
- McKay, R., & Furl, N. (2017). Doppelgängers and dissociations: Lesion network mapping illuminates misidentification delusions. *Brain*, 140(2), 262-265. doi: 10.1093/brain/aww323
- Miyazono, K., Bortolotti, L., & Broome, M. R. (2015). Prediction-error and two-factor theories of delusion formation: competitors or allies? In N. Galbraith (Ed.), *Aberrant beliefs and reasoning*. Hove and New York: Psychology Press.
- Nie, C. (2016). Delusional beliefs, two-factor theories, and bizarreness. *Frontiers of Philosophy in China*, 11(2), 263-278. doi: 10.3868/s030-005-016-0020-1