Philosophers' Imprint

VOLUME 25, NO. 18
JULY 2025

ENCAPSULATED FAILURES

Zoe Jenkin

Rutgers University

doi.org/10.3998/phimp.3809

1. Introduction

Homer's *Odyssey* teaches many lessons, but one of its greatest is that we should always respond to our reasons. As Odysseus and his crew sail away from the island of the Cyclops, having just blinded the one-eyed giant with a stake, Odysseus feels the pull of a strong reason to keep his identity a secret. He knows that while the Cyclops is temporarily impotent, towering on the shore in blind rage, his father is the sea god Poseidon who has the power to inflict a vengeful wrath. Yet Odysseus is overcome by pride at his own cleverness and shouts his own name from the prow of his ship, carelessly jeopardizing the safety of his crew.

The parable of Odysseus and the Cyclops is a uniquely rich and compelling story, but it involves an utterly ordinary kind of failure to respond to reasons: ego eclipses prudence. In this moment, Odysseus is irrational and is responsible for this irrationality. His irrationality stems from the fact that he has violated the rational requirement to respond to his reasons. While we often meet this requirement in everyday life, we also often violate it by failing to respond to reasons due to ego, closed-mindedness, carelessness, or other poor epistemic habits. In such cases, our failures render us irrational.

My focus here is on how the requirement to respond to reasons applies to beliefs. On ideal epistemologies that abstract from the realities of human cognition, the requirement to respond to reasons is universal—all beliefs are required to respond to reasons. However, on non-ideal epistemologies that are sensitive to the realities of human cognition, the scope of the requirement to respond to reasons is complicated by our cognitive architecture.[1] The human mind is not one perfectly unified system but instead breaks down into informationally encapsulated fragments. These fragments occur throughout belief, language processing,

1. For discussion of ideal and non-ideal epistemologies, see Pasnau (2013), Staffel (2017), and Carr (2021).

perception, and emotion.[2] Informational encapsulation can prevent a belief stored in one part of the mind from responding to a reason stored in another part of the mind. For example, even if you know your friend is furrowing their brow because they have a headache, after glancing at them you might automatically believe that they are angry. Your failure to respond to the reason that they have a headache is not due to your individual epistemic character but rather to the informational encapsulation of emotion detection (Smortchkova, 2017). I call such failures to respond to reasons due to informational encapsulation 'encapsulated failures.'[3]

This paper considers how informational encapsulation impacts and constrains the scope of the requirement to respond to reasons. Are beliefs always required to respond to reasons, even in cases of informational encapsulation? Or is informational encapsulation outside the scope of the requirement to respond to reasons? This question is important because the scope of the requirement to respond to reasons determines the scope of our irrationality for its violation.

Several related considerations put pressure on the idea that our beliefs are required to respond to reasons in cases of informational encapsulation. First, the informationally encapsulated structures of cognitive architecture are not freely chosen. They are either evolved or formed unintentionally through our patterns of thought and action. Second, from a subjective perspective, it is not obvious how we could respond to reasons when informational encapsulation prevents us from doing so. This violates the principle that "ought implies can" (Kant, 1787/1997). Third, attitudes that are not directly sensitive to our reasons are often thought to not be truly ours (Moran, 2001; Boyle, 2009). If an attitude is not ours, it cannot make us irrational. These considerations suggest that beliefs produced by encapsulated systems are exempt from the requirement to respond to reasons. I call this position 'Lenience.'

On the other hand, cognitive architecture is an internal feature of an agent's mind, not an external coercive force of the kind that typically absolves one of responsibility (Frankfurt, 1973; Scanlon, 2015). Cognitive architecture shapes the most basic aspects of our mental lives and so is deeply interwoven with who we are. Such a central feature of our rational agency should be reflected in rational evaluation. These considerations point toward a position on which the requirement to respond to reasons applies to beliefs produced by encapsulated systems just as it does to beliefs produced by unencapsulated parts of the mind.[4] Failures to respond to reasons are equally irrational whether they are caused by informational encapsulation, laziness, or ego. I call this position 'Austerity.'

Depending on which of the above considerations one emphasizes, cognitive architecture seems to vacillate between being a fundamental part of who we are and a constraint that limits our agency. How should we adjudicate between these two polarized positions? My strategy is

2. I do not hold that all the above listed parts of the mind are perfectly informationally encapsulated. As I discuss in §2, informational encapsulation comes in degrees. For example, while vision is significantly encapsulated from belief, it is nonetheless diachronically penetrable through perceptual learning (Goldstone, 1998). The examples I use throughout the paper support the view that the mind contains at least several systems that are informationally encapsulated to a substantial degree. For additional defenses of informational encapsulation and fragmentation, see Fodor (1983), Cherniak (1986), Pylyshyn (1999), Egan (2008), Quilty-Dunn & Mandelbaum (2018), Bendaña & Mandelbaum (2021), and Elga & Rayo (2021).

3. There are two types of encapsulated failures: In the first type, a belief fails to respond to a reason because the belief is produced by an informationally encapsulated system that cannot access central cognition (or another mental subsystem) in which the reason is stored. In the second type, a belief fails to respond to a reason because the belief is produced by central cognition, which cannot access the encapsulated system in which the reason is stored. Fodor (1983) refers to the second type as 'limited central accessibility." My focus in this paper is on the first type of case, but the latter likely admits of similar treatment.

4. Borgoni (2021) argues for this kind of view. Proponents of global coherence requirements on rationality hold the related view that coherence is required across encapsulated parts of the mind (e.g., Davidson, 1982, 1983; Stalnaker, 1984). Proponents of only local coherence requirements hold that coherence is only required within fragments—a position similar to Lenience (e.g., Cherniak, 1983, 1986; Yalcin, 2021; Elga & Rayo, 2022). My focus is on reasons-responsiveness requirements rather than coherence requirements, but I return to the topic of coherence in §3 and §4.

to examine a broad spectrum of types of encapsulated failures and our corresponding judgments of rationality or irrationality. I argue that neither Lenience nor Austerity adequately explains our intuitive judgments of rationality and irrationality across the full spectrum of encapsulation. I argue that a third view called 'Architectural Sensitivity' does much better at this task, while also accommodating the motivating insights of both Lenience and Austerity. Architectural Sensitivity is the view that beliefs produced by encapsulated systems are required to respond to reasons, but when they fail to do so their degree of irrationality is mitigated by informational encapsulation. This view respects the idea that our human limitations, such as informational encapsulation, are at once fundamental rational flaws and individually absolving.

The rest of the paper proceeds as follows. In §2, I unpack the requirement to respond to reasons and its role in rational evaluation. In §3, I introduce a spectrum of encapsulated failures. Through the felt irrationality of these cases, I argue against Lenience. In §4, I argue against Austerity by highlighting disanalogies in our intuitions about the rationality of encapsulated and unencapsulated failures. In §5, I introduce Architectural Sensitivity and argue that this view better accounts for our intuitions about the spectrum of encapsulation. In §6, I argue that Architectural Sensitivity does not reduce to a control requirement.

2. The Requirement to Respond to Reasons

What exactly is the requirement to respond to reasons? We have already seen a version of it at work in Odysseus's misguided deliverance from the Cyclops. Here is the requirement as it applies to beliefs:

> **Respond to Reasons:** An agent's beliefs are rationally required to respond to all the agent's relevant reasons.

This requirement has deep roots in epistemology. It resembles what Van Inwagen refers to as 'Clifford's Other Principle': "It is wrong always, everywhere, and for anyone to ignore evidence that is relevant to his beliefs, or to dismiss relevant evidence in a facile way" (Van Inwagen, 1996, 145).[5] The requirement to respond to reasons is also implicit in many of our everyday rational assessments, for example when we criticize someone for neglecting an important consideration in their belief-formation.

Adherence to rational requirements determines a belief's rational status. Thus, Respond to Reasons comes with a corollary:

> **Irrationality:** If an agent's belief fails to respond to all the agent's relevant reasons, that belief is thereby irrational to at least some degree.

Irrationality is a sufficient condition on a belief's irrationality (although I note some exceptions later in this section). I leave open whether Respond to Reasons is a sufficient condition on a belief's rationality. Rationality may also require adherence to norms of inquiry (Smith, 2014; Goldberg, 2017; Friedman, 2020), proper allocation of attention (Watzl, 2017), coherence, or more particular forms of response (e.g., Bayesian updating). Nonetheless, I take responding to reasons to be a significant part of what makes beliefs rational.

Reasons are considerations in favor of acting, believing, or mentally representing in a certain way (Scanlon, 2014).[6] My arguments are neutral as to whether reasons are ultimately facts (e.g., Williamson, 2000), propositions (e.g., Fantl & McGrath, 2009) or mental states (e.g., Davidson, 1963). Whatever the essential nature of reasons, agents possess reasons in virtue of their mental states. If reasons are facts or propositions, mental states epistemically relate agents to those facts or propositions. For example, Odysseus's visual experience as of the cyclops waving his fists on the shore epistemically relates him to the

5. See also Clifford (1877) and Chignell (2018). For broader discussion of the requirement to respond to reasons, see Conee & Feldman (1985), Williamson (2000), Kelly (2006), Worsnip (2018), and Gertler (2021).
6. I include 'mentally representing' here to allow for the possibility that reasons can also be considerations in favor of e.g., perceiving the world in a certain way, or having certain emotions.

fact or proposition that the cyclops is angry. I use the terminology of mental states 'providing reasons' to the agent, which should be read as compatible with factualist, propositionalist, and mentalist accounts of reasons.

Respond to Reasons only applies to the reasons an agent possesses. It concerns one's own internal network of mental states and possessed reasons.[7] It is as a sufficient condition on possessing a reason that an agent can use that reason to guide their inferences and/or actions. I do not claim that agents possess reasons provided by mental states that are stored solely within an encapsulated system and never used beyond that system. My examples in §3-§5 of failures to respond to reasons involve reasons that the agent can use to guide actions or inferences in other parts of the mind.

The requirement to respond to reasons focuses on agents' *relevant reasons*. At any given time, an agent possesses a massive number of reasons, far more than those to which they could feasibly respond. Only a small subset of those reasons is relevant to any given belief. For example, my possessed reason that France borders Switzerland is irrelevant to my belief that cheddar is a kind of cheese but is relevant to my belief that it is possible to drive from Paris to Bern. This is because the proximity of France to Switzerland bears on the truth of the latter belief, but not the former.[8] If an agent fails to respond to the reason that France borders Switzerland in forming the belief that cheddar is a kind of cheese, they are not in violation of the requirement to respond to reasons. In general, an agent's relevant reasons are their reasons that bear on the truth of the belief in question, at least as far as epistemic reasons are concerned.[9]

On evidentialist views, only epistemic reasons are relevant to belief (e.g., Clifford, 1877; Moran, 1988; Shah, 2006). This makes bearing on the truth of a belief the end of the story for relevance of reasons. However, on pragmatist views, factors beyond bearing on truth can make reasons relevant to belief (e.g., James, 1896; Foley, 1992; Rinard, 2018). For pragmatists, prudential and/or moral reasons can be relevant to a belief because they make that belief attractive (or unattractive) to an agent in light of their desires and ends. To take a classic example, some pragmatists take the reason that believing in God may save you from eternal damnation to be relevant to your belief in God, whereas evidentialists do not (James, 1896).

The requirement to respond to reasons will be applied differently depending on whether one is an evidentialist or pragmatist, as well as depending on one's particular version of evidentialism or pragmatism. For at least some pragmatists, an agent can violate the requirement to respond to reasons by failing to respond to a prudential or moral consideration that makes a belief attractive or unattractive to her.[10] Evidentialists would not count this as a violation. Adjudicating between evidentialism and pragmatism is beyond the scope of this paper, so going forward I focus on examples of failures to respond to truth-relevant, epistemic reasons that count as violations of the requirement for evidentialists and pragmatists alike.

The requirement to respond to reasons says that beliefs are required to respond to reasons, but it leaves open exactly what kind of response is required. The amount of response required might be perfect evidential integration or an incremental push in the direction of the reason's force. Some reasons are merely pro tanto, so even if a belief were to respond

7. Other epistemic norms (e.g., norms of inquiry) may dictate what kind of reasons agents are required to bring into their possession. For discussion see Smith (2014), Goldberg, (2017), and Friedman (2020).
8. Perhaps in some sense everything bears on the truth of everything else, a la a Quinean web of belief (Quine, 1951). I set that issue aside here. I have in mind a more commonsense notion of 'bearing on' according to which only some things bear on other things.
9. Relevance has also been theorized in the contexts of linguistic communication (e.g., Sperber & Wilson, 1995) and inquiry (e.g., Freidman, 2020). While some insights from these contexts may bear on beliefs' responses to reasons, I set them aside in the interest of space.
10. I say 'for at least some pragmatists' here because there are many subtle variations of pragmatism. On some versions, responding to practical and/or moral reasons is permitted but not required.

to them perfectly, they would be outweighed by other considerations. Responding to reasons might result in adjusting a belief, forming a new belief, or eliminating a preexisting belief.

My focus here is on the application of the requirement to respond to reasons to beliefs, but I also discuss potential applications to attitudes beyond belief, such as perception and emotion, in §3. Some version of the requirement plausibly applies to agents as well. Not only beliefs, but also agents are under rational pressure to respond to their possessed reasons. However, the precise nature of this rational pressure may be more complex than the simple agential equivalent of belief's requirement to respond to reasons, according to which agents are required to respond to all their possessed reasons in order achieve rationality. This is because agential rationality may be comprised not only of an agent's responses to reasons, but also of their capacity to cultivate and monitor their reason-giving mechanisms (e.g., Jones, 2003; Carman, 2018).[11] So while the rationality of beliefs and agents is deeply intertwined, there may not always be a neat one-to-one correspondence. For this reason, I focus on the rationality of individual attitudes rather than whole agents.

What is the scope of belief's requirement to respond to reasons? Our epistemic practice suggests the requirement to respond to reasons applies to most of our ordinary beliefs. But some are outside its scope. For example, a belief is not required to respond to a reason that has too many conjuncts for any human mind to grasp. Or a belief is not required to respond to a reason if an agent's attention is wholly devoted to more important matters. These exemptions are due to basic features of cognitive architecture: capacity limits on memory and attention.[12]

Cognitive architecture is the structure of the mind. It includes a basic system of information processing, such as computations on structured symbolic representations (Fodor, 1976; Pylyshyn, 1984) or probabilistic Bayesian inferences (Griffiths, Kemp, & Tenenbaum, 2008). It also

11. I thank an anonymous reviewer for raising this point.
12. There may be other exemptions from the requirement to respond to reasons that are not due to cognitive architecture, e.g., when a reason concerns a wholly unimportant matter.

includes the way the mind breaks down into parts, such as memory stores, the language faculty, and the different perceptual subsystems, as well as how those parts interrelate. Cognitive architecture need not be innate—it can also include structural divisions or mechanisms that are learned from experience.

If some aspects of cognitive architecture limit the scope of the requirement to respond to reasons, it is natural to inquire as to whether other aspects of cognitive architecture do the same. My focus here is on informational encapsulation. Informational encapsulation is a central feature of cognitive architecture because it helps delimit the boundaries of mental systems. A system is informationally encapsulated if it can only access information stored within that system, and not information stored in other parts of the mind (Fodor, 1983). For example, vision displays informational encapsulation through the persistence of illusions. The visual system cannot access our beliefs debunking the illusion, so visual processing proceeds unaffected and the illusion persists.

Informational encapsulation is most widely discussed as a feature of modular systems such as perception and the language faculty. Modular systems also share other features, including domain specificity, automaticity, fast processing, shallow outputs, limited central accessibility, and characteristic patterns of development and breakdown (Fodor, 1983). Informational encapsulation plays a central role in delineating the boundaries of modules, but an informationally encapsulated system can lack the other features of modularity. One place this occurs is belief, where informationally encapsulated pockets of beliefs are often described as fragments rather than modules (e.g., Lewis, 1982; Egan, 2008; Bendaña & Mandelbaum, 2021; Borgoni, 2021; Elga & Rayo, 2021; Gertler, 2021; Yalcin; 2021). I speak in terms of informational encapsulation rather than fragmentation to highlight the presence of an informational boundary, but both terms describe the phenomenon under consideration.

Informational encapsulation comes in degrees. Degrees of informational encapsulation are modulated by several factors, including which information can be accessed, frequency of information access, speed

of information access, the range of contexts in which information can be accessed, the size of the drain on resources when information is accessed, and the parts of the mind that are accessible. For example, the language faculty regularly accesses information from vision and audition, but not from olfaction, gustation, or tactition. In contrast, belief formation regularly accesses information across all sensory modalities. On the opposite end of the spectrum, visual edge detection relies on visual contrast cues, and rarely, if ever, takes inputs from other sensory modalities (Marr, 1982; Georgeson et al., 2007). Thus, the language faculty is more encapsulated than belief formation, but less encapsulated than edge detection (at least along the dimension of parts of the mind that are accessible).

3. Against Lenience

The first position on the rationality of informational encapsulation that I will consider is called 'Lenience:'

> **Lenience:** Beliefs are not required to respond to reasons in cases of informational encapsulation.

According to Lenience, encapsulated failures are entirely outside the scope of the requirement to respond to reasons. It follows that beliefs are not made irrational by encapsulated failures.

Why might this be so? Two theoretical notions underpin Lenience: volitionalism and 'ought' implies 'can'. First, volitionalism is the idea that we are only responsible for what we can choose or control.[13] Encapsulated failures are not under our control (or so the argument goes), so according to volitionalism they are not our responsibility. A related idea is that beliefs that do not respond to our reasons are not truly our own, and thus not our responsibility (Moran 2001, Boyle 2009). Second, 'ought' implies 'can' is the Kantian idea that norms can only demand actions we are able to perform (Kant, 1787/1997; Vranas, 2007). Informational encapsulation renders beliefs unable to respond to reasons, so according to 'ought' implies 'can,' such responses cannot be required.[14]

While Lenience is both intuitively and theoretically appealing, it does not align with many of our judgments about individual cases. In this section, I consider a spectrum of encapsulated failures across diverse parts of the mind. This spectrum includes failures to respond to reasons due to belief fragmentation, emotion, and automatic perceptual belief. These systems lead to encapsulated failures that are intuitively irrational, indicating that Lenience is false.

Consider first belief fragmentation (Egan, 2008; Bendaña & Mandelbaum, 2021; Elga & Rayo, 2021). In belief fragmentation, pockets of beliefs become informationally isolated from each other, leading to inconsistencies. To illustrate, consider a classic example from David Lewis:

> I speak from experience as the repository of a mildly inconsistent corpus. I used to think that Nassau Street ran roughly east-west; that the railroad nearby ran roughly north-south; and that the two were roughly parallel. . . Once the fragmentation was healed, straightway my beliefs changed: now I think that Nassau Street and the railroad both run roughly northeast-southwest. (Lewis, 1982, p. 436)

Lewis's beliefs are resistant to integration due to their storage in separate, informationally encapsulated fragments. These fragments are likely formed due to acquiring information in different contexts. Perhaps Lewis acquired his belief that Nassau Street runs east-west

13. Volitionalism is most widely discussed in the domain of moral responsibility (e.g., Fischer & Ravizza, 1998; Levy, 2005; Wolf, 1990).

14. Volitionalism and 'ought' implies 'can' also challenge the idea that there are any rational requirements on belief at all, given arguments for doxastic involuntarism (the view that belief is not under our voluntary control (Alston, 1988)). Several responses to this challenge are available, e.g., the argument that at least some beliefs are under the requisite kind of control (Audi, 2008; Weatherson, 2008; Wedgewood, 2013) or the rejection of the epistemic 'ought' implies 'can' (Feldman, 2003). I assume here that some such response is successful and at least some beliefs are subject to rational requirements.

by walking along Nassau Street to get to the philosophy department, which he knew to be on the east side of campus. Perhaps he acquired his belief that the train tracks run north-south by consulting a map. These fragments remain isolated so long as they are only used in separate contexts (walking and map-reading). We might find ourselves with similar kinds of belief fragments for professional vs. social contexts, political vs. personal decisions, or abstract vs. concrete reasoning tasks.

Lewis's inconsistent corpus is a paradigm of irrationality. The natural explanation for this irrationality is that his beliefs have violated the requirement to respond to reasons. Lewis possesses all the relevant reasons to support the conclusion that Nassau Street and the railroad run northeast-southwest (e.g., memories of walks, the locations of landmarks, the directions of other roads) and can use these reasons to guide his inferences and actions in other contexts. But in the context described above, he fails to do so. Lewis's beliefs' irrationality indicates that such ordinary cases of belief fragmentation are within the scope of the requirement to respond to reasons.

One might wonder whether this irrationality could instead be explained by Lewis's violation of a coherence requirement according to which beliefs must cohere on pain of irrationality.[15] I grant that coherence requirements may play a role in explaining why Lewis's beliefs seem irrational. However, there is an aspect of his beliefs' irrationality that stems specifically from their failure to respond to reasons. Consider a scenario in which Lewis randomly revises his directional beliefs on a whim, irrespective of reasons, and luckily ends up with the belief that Nassau Street and the railroad tracks both run northeast-southwest. While Lewis has improved his beliefs' epistemic situation with respect to coherence and truth, there is still something rationally amiss. This lingering irrationality is explained by his violation of the requirement to respond to reasons. His beliefs cohere, but he has still neglected his own reasons in forming them.[16]

In the passage quoted above, Lewis eventually heals his fragmentation and returns to a state of rationality. While his initial fragmentation is due to the baseline architectural tendency of human minds to fragment, the boundaries of his fragments are not indelibly fixed. By directing his attention to his inconsistency, he can either dissolve the fragments or siphon information from one to the other. In this way Lewis's belief fragments are unlike perception and language processing, whose informational borders are not so easily crossed. Thus, one might think the requirement to respond to reasons only applies here because belief fragments do in fact respond to reasons as soon as we deploy sufficient cognitive effort.

Yet irrationality also arises in systems whose fragments are not so easily dissolved. Consider visual perception, which is significantly encapsulated from cognition (Fodor, 1983; Pylyshyn, 1999).[17] Evidence for perception's encapsulation comes from the persistence of illusions. For example, you will perceive a pencil in a glass of water as bent, even if you know everything about the physics of light refraction. Additionally, each time you glance at the pencil you will continue to have the automatic (if fleeting) perceptual belief that the pencil is bent,

15. For discussion of coherence requirements, see e.g., Bonjour (1985), Lehrer (2000), Berker (2015), and Worsnip (2018).

16. One might wonder whether this rational difference can be explained Lewis's lack of justification rather than by his violation of the requirement to respond to reasons. We can see that this is not the case because even if Lewis resolves his incoherence due to an independent source of justification, such as testimony, the intuition that Lewis's directional beliefs are irrational to at least some degree irrational remains. Lewis's beliefs might be justified, but they are still rationally flawed due to their insensitivity to his prior reasons.

17. I do not claim that perception is perfectly encapsulated. Because informational encapsulation is a degreed notion, my claim that perception is informationally encapsulated is compatible with various forms of cognitive influence on perception (e.g., Prinz, 2006; Macpherson, 2012; Lupyan, 2015).

despite knowing that it truly is not.[18] When you reach for the pencil, it will be hard to resist responding to this perceptual belief by grabbing at the as-if-bent location. Perceptual beliefs are formed automatically in response to perception, relying on a proprietary information database.[19] This makes perceptual belief formation fast and efficient. It also prevents perceptual beliefs from responding to reasons provided by beliefs in central cognition. I focus on perceptual beliefs rather than on perception itself so as not to assume that perception is rationally evaluable. If perception is rationally evaluable (Siegel, 2017; Jenkin, 2023a), then the question of whether encapsulated systems are required to respond to reasons is even more pressing due to the high degree of encapsulation in perception.

While the automatic perceptual belief that the pencil in water is bent does not seem obviously irrational, examples from perceptual learning illustrate that the requirement to respond to reasons does sometimes extend to perceptual beliefs. While vision is synchronously encapsulated, it can be influenced by beliefs through diachronic learning (Goldstone, 1998), which is a way of responding to reasons over time (Jenkin, 2023a). Consider elite chess masters who learn to visually identify available moves through a perceptual learning process called unitization (Chase & Simon, 1972; Charness et al., 2001; Campitelli et al., 2007; Bilalic et al., 2010). Through years of experience combining their visual representations of chessboards and their encyclopedic beliefs about the rules of chess, their visual systems store information about which configurations of pieces make up salient units. Their visual systems apply this stored information to raw visual stimuli, producing perceptual representations of the chess board as segmenting into available moves. This unitization aids memory of the board and facilitates game strategy.

When a chess master sees the board and automatically believes there is a castling, this belief is a diachronic response to reasons provided by their beliefs about the rules of chess. For example, when a chess master looks at a board with a rook and a king in locations x and y, their belief that a castling is available is a response to their reason (provided by a belief) that when a rook and king are in positions x and y, a castling is available. Their beliefs about the rules of chess diachronically influences their visual system as they learn to see castlings and automatically believe they are available. The chess master's belief that there is a castling seems not only useful and accurate, but also rational. A good explanation of this rationality is that this belief meets the requirement to respond to reasons through perceptual learning. It is particularly important to have a good explanation of this belief's rationality, because such beliefs constitute a major part of chess expertise.

I grant that there may be other explanations of an attitude's rationality aside from meeting the requirement to respond to reasons. For example, one might think the chess master's belief that there is a castling is rational because it coheres with their beliefs about the rules of chess. However, this coherence does not explain an important rational contrast: the chess master's belief in the castling seems *more* rational than the belief of a beginner who sees the castling due to luck rather than due to a response to their reasons. Yet both beliefs are equally coherent. The chess master's rationality stems from their response to reasons.

One might also wonder whether the chess master's belief is rational because it is supererogatory, rather than because it meets the require-

18. This attitude is a belief both according to views on which belief is a relation to a representation with a particular psychofunctional profile (e.g., Quilty-Dunn & Mandelbaum, 2018) and on which belief is a set of dispositions (e.g., Schwitzgebel, 2002). It is also a belief according to views on which beliefs must be revisable in light of evidence (e.g., Helton, 2020) because even if perceptual beliefs are automatically formed, they can be revised. If one thinks belief requires an agent's commitment to the truth of a proposition, one might deny that this attitude is a belief and instead classify it as a mere thought. If so, this example can be understood as about the encapsulation of automatic perceptual thought rather than belief.

19. See Gilbert (1991), Mandelbaum (2013) and Quilty-Dunn (2015) for arguments that perceptual beliefs are informationally encapsulated. See Fodor (1983) and Lyons (2011) for arguments that perceptual beliefs are at least sometimes influenced by central cognition. My arguments here only require that perceptual beliefs are significantly encapsulated, not perfectly encapsulated, so my arguments are compatible with both positions.

ment to respond to reason.[20] Perhaps the chess master is the epistemic equivalent of someone who donates blood to the Red Cross every 56 days (as often as permitted in the United States)—that is, a person deserving of praise precisely because they go beyond the call of duty. If so, the requirement to respond to reasons need not apply.

While this is a plausible diagnosis of the chess master in isolation, nearby examples illustrate that supererogation cannot be the end of the story. Consider an amateur chess player who is cognitively resistant to integrating their knowledge of chess with their visual system. They know the rules of chess and have been playing for the length of time that is typically sufficient for unitization. They possess the rules of chess as a reason, as evinced by their ability to explain them to other players. Yet their visual system fails to respond to their knowledge of the rules of chess by storing information about how to identify available moves. This manifests in their visual experience: when they scan a board with a castling, they do not perceive it, even though they know in the abstract that when a rook and king are in their current positions, a castling is available. In response to this visual experience, the amateur player automatically forms the false perceptual belief that there is no castling and chooses a different move. This belief seems irrational because they possess all the reasons that they need to identify the castling, and they have ample time to respond to those reasons through perception, yet they fail to do so.[21] This irrationality is well-explained by the view that encapsulated perceptual beliefs are governed by the requirement to respond to reasons, not merely by supererogation.[22]

Recalcitrant emotions provide further support for the view that the requirement to respond to reasons extends to encapsulated systems. Emotional processing is encapsulated from much of central cognition.[23] Consider imposter phenomenon and the family of related workplace and academic anxieties. These experiences are characterized by feelings of fear, pessimism, unworthiness, and low self-esteem, even when one knows they are irrational.[24] These feelings persist even when one possesses good reasons that one is qualified and well-prepared, such as academic awards, promotions, praise from colleagues, and journal acceptances. The fact that these emotions cannot be intellectualized away reflects that emotion is informationally encapsulated from the reasons our beliefs provide.

Cognitivists and non-cognitivists about emotion agree that recalcitrant emotions are paradigms of irrationality (D'Arms & Jacobson, 2003; Brady, 2009; cf. Hursthouse, 1991). This irrationality indicates that like beliefs, emotions are subject to the requirement to respond to reasons. The irrationality of recalcitrant emotions stems from their violation of this requirement. Informational encapsulation does not exempt emotions from the requirement to respond to reasons.

The examples of recalcitrant emotions and perceptual learning discussed above raise the question of whether there is a timeframe on reasons-responsiveness. Does a belief or emotion only count as respon-

20. I thank Allan Hazlett for suggesting this possibility.

21. This is plausibly an example of a mismatch between propositional knowledge and skill: the advanced amateur has propositional knowledge of the rules of chess but lacks the skill of perceptually identifying available moves. This analysis is compatible with, and even helps explain, the idea that the advanced amateur's belief that there is no castling is unjustified: the belief is generated by a system that is deficient in a skill it was well-equipped to acquire but failed to do so. For further discussion of the epistemology of skill, see Pavese (2025). For further discussion of responding to reasons through perceptual learning, see (Jenkin, 2023a). Thanks to an anonymous reviewer for raising this issue.

22. One might wonder where this leaves true beginners who fail to unitize the board because they lack both knowledge and experience. True beginners skirt the requirement to respond to reasons because they do not possess any relevant reasons (i.e., the rules of chess, and specifically the rule about what constitutes a castling). This explains why beginners do not seem irrational for failing to identify available moves.

23. The informational encapsulation of emotion is less straightforward than that of the other systems discussed here because emotional processing is largely driven by valence. Nonetheless, emotions lack access to information stored in other parts of the mind, which defines encapsulation. For discussion of the informational encapsulation of emotion, see de Sousa (1987), Cosmides & Tooby (2000), Prinz (2004), and Majeed (2019, 2020).

24. For an overview of imposter phenomenon, see Clance & Imes (1978). For an argument that imposter phenomenon is in fact rational, see Slank (2019).

sive to a given reason if it responds within a certain amount of time? As I understand it, there is no such strict time limit. Monitoring and adjusting our emotions over long periods of time is still a way of responding to reasons (Jones, 2003; Carman, 2018). However, a slower response to reasons is often a way of being less responsive to reasons, at least along a certain dimension. A slower response to reasons is also often a way of being more informationally encapsulated. I return to this issue in §5.

This section has detailed how failures to respond to reasons across different encapsulated systems are irrational, despite being caused by informational encapsulation. These examples show that Lenience, the view that the requirement to respond to reasons does not apply to encapsulated systems, is incorrect. Lenience not only doles out too many free passes but also denies rational credit where it is due, as in the case of the chess master. Together, these examples build a strong case that the scope of the requirement to respond to reasons includes encapsulated systems.

4. Against Austerity

The natural alternative to Lenience is Austerity:

> **Austerity**: Beliefs are required to respond to an agent's relevant reasons in cases of encapsulation, just as they are absent encapsulation.

According to Austerity, encapsulated failures are within the scope of the requirement to respond to reasons. It follows that when beliefs fail to respond to agents' relevant reasons, they are irrational, irrespective of encapsulation.

Why endorse Austerity? Austerity is motivated by the idea that our cognitive architecture is part of who we are as rational agents, so it should redound on our rational statuses. Given the arguments against Lenience in the previous section, Austerity looks particularly tempting. However, I will argue that Austerity does not fare much better when held up to the spectrum of encapsulated failures.

The first problem with Austerity is that it delivers verdicts that are simply too harsh. Consider again the classic example of known illusions. You look at a stick in water and perceive it as bent, and then automatically believe it is bent, despite knowing that the illusion is caused by light refraction. Your automatic perceptual belief does not seem irrational, even though it fails to respond to your reason. This perceptual belief seems exempt from the requirement to respond to reasons precisely because of its encapsulation.

Austerity also fails to capture important rational differences between encapsulated and unencapsulated systems. For example, recalcitrant emotions seem less irrational than analogous recalcitrant beliefs. Consider two employees who experience imposter phenomenon in different ways. They both have strong reasons to believe they are performing well at work—positive progress reviews, compliments from colleagues, raises, and so on. Employee A has their beliefs in line with these reasons, but their emotions are recalcitrant. They are plagued by irrational feelings of anxiety and inadequacy. Employee B has their emotions in line with these reasons, feeling confident and relaxed, but they persistently believe that they will be fired at any moment. While both employees' mental states seem irrational to a degree, B's beliefs seem more irrational than A's emotions. This disparity can be explained by the idea that the requirement to respond to reasons does not apply equally to encapsulated and unencapsulated systems.

Might there be some alternative explanation for why B's recalcitrant beliefs seem more irrational than A's recalcitrant emotions? One possibility is that these agents possess relevant reasons for belief (e.g., positive performance reviews, compliments from colleagues, raises, etc.), but not relevant reasons for emotion.[25] Because emotion and belief are different attitudes, the reasons-for relation may differ between them.[26] If so, possession of reasons for belief does not guarantee possession of reasons for emotion. Do these agents truly possess reasons to *feel*

25. I thank an anonymous reviewer for suggesting this possibility.
26. I set aside views on which there are no reasons for emotions, e.g. (Maguire, 2018; Naar, 2022; Schultz, 2025). For other views on which there are reasons for emotions, see e.g., Brady (2013) and D'Arms & Jacobson (2000).

confident in their work performance?

Initial support for a 'yes' answer comes from our ordinary ways of speaking and thinking about emotion. When an employee receives praise from their supervisor, it is natural to say that they have good reason to feel confident. If they fail to feel confident (like employee A), it is natural to say that their feelings are irrational precisely because they have good reason to feel otherwise.

Further support for a 'yes' answer comes from consideration of the conditions on possessing reasons. As noted earlier, an individual possesses a reason if they can use that reason to guide inferences or actions in some part of their mind. But the conditions on possessing reasons for particular types of mental attitudes, such as beliefs or emotions, might be more stringent. For example, possessing a reason for belief might require that the agent be able to use that reason to guide their beliefs. Possessing a reason for emotion might require that the agent be able to use that reason to guide their emotions.

Even given these more stringent condition on possessing reasons, both employees possess reasons for emotions. Employee B (who has recalcitrant beliefs) *does* use the praise from their supervisor to guide their emotions (resulting in a feeling of confidence) and so clearly *can* do so. Employee A (who has recalcitrant emotions) does not use the praise from their supervisor to guide their feeling of self-confidence, but they nonetheless can do so, at least over time and with the assistance of psychotherapy or other cognitive exercises. Emotions are by and large diachronically malleable, even when they are synchronically resistant. Additionally, Employee A can use the reason of their supervisor's praise to guide other emotions, such as warm feelings toward the supervisor for their support. Thus, the reasons that are neglected in cases of recalcitrant emotions are nonetheless possessed.

A second alternative explanation for the rational disparity between A and B is that they fare differently with respect to coherence requirements. Consider the broad coherence requirement that says our beliefs and emotions must cohere on pain of irrationality. A's belief that they are performing well does not cohere with their anxious emotions, and B's belief that they will be fired does not cohere with their confident emotions. Both A and B fail to meet this coherence requirement and are thereby irrational. This does not explain why B's beliefs seems more irrational than A's emotions.

One might suggest that the coherence requirement applies only to beliefs, thus explaining why B's beliefs seems more irrational than A's beliefs (because B's beliefs do not cohere whereas A's do). However, this fails to explain why A's emotions still seem irrational to a degree, just less so than B's beliefs. While coherence may play some role in explaining the irrationality of recalcitrant emotions, it is not the end of the story.[27]

5. Architectural Sensitivity

The spectrum of encapsulation illustrates that both Lenience and Austerity are unsatisfactory. In this section, I propose an alternative that better captures our intuitive verdicts about the spectrum of encapsulation. I call this alternative Architectural Sensitivity.

To understand Architectural Sensitivity, first recall Respond to Reasons and Irrationality:

> **Respond to Reasons**: An agent's beliefs are rationally required to respond to all the agent's relevant reasons.
> **Irrationality:** If an agent's belief fails to respond to all the agent's relevant reasons, that belief is thereby irrational to at least some degree.

27. One might suggest that the coherence requirement applies to both encapsulated and unencapsulated systems, but with unequal force. This is roughly a coherence-based version of my own proposal in §5. I am open to such a coherence requirement, but the requirement to respond to reasons is nonetheless needed to explain the intuitive rational difference between equally coherent belief systems that differ as to whether their coherence is due to responding to reasons or luck (e.g., the chess master and the lucky beginner discussed in §3).

Architectural Sensitivity modulates Irrationality:[28]

Architectural Sensitivity: The degree of irrationality of an agent's belief that fails to respond to all the agent's relevant reasons due to informational encapsulation is inversely proportional to the degree of encapsulation of the subsystem that produced the belief.

A first thing to note about Architectural Sensitivity is that it applies specifically to beliefs that fail to respond to reasons *due to informational encapsulation*. It does not apply to all states that fail to respond to reasons and are produced by informationally encapsulated systems. Consider a state produced by reasoning within a belief fragment that fails to respond to a reason stored within that very belief fragment due to indolence. The agent has perfectly good access to this reason, but they are too lazy to consider it. Such cases are not within the purview of Architectural Sensitivity. Architectural Sensitivity only applies when informational encapsulation prevents a belief from responding to a reason.

In such cases, Architectural Sensitivity says that the belief's degree of irrationality is inversely proportional to the degree of encapsulation of the subsystem that produced it. If the subsystem is unencapsulated, the belief is highly irrational. If the subsystem is weakly encapsulated, the belief is significantly irrational, but slightly less so (all else held equal). If the subsystem is moderately encapsulated, the belief is moderately irrational. If the subsystem is strongly encapsulated, the belief is weakly irrational. If the subsystem is perfectly encapsulated, the belief is not irrational. Perfectly encapsulated systems occur rarely if ever, so failure to respond to reasons almost always leads to irrationality of some degree.[29]

The idea of a system's degree of encapsulation is central to Architectural Sensitivity. As discussed in §2, several factors determine a system's degree of encapsulation, including what parts of the mind can be accessed, what domains of information can be accessed, frequency of information access, speed of information access, the contexts in which information can be accessed, and the size of the drain on resources that information access causes. The sum of these factors determines the degree of encapsulation of a subsystem, and hence the degree to irrationality of its encapsulated failures.

The idea of degree of (ir)rationality is also central to Architectural Sensitivity. This is the commonsense notion that rationality is not an on/off switch. There are ways of being more or less rational or irrational. For example, a belief formed through a multi-step inference is more irrational if that inference involves jumping to conclusions twice than if it involves jumping to conclusions only once. Mental states are made more or less rational by how well they respond to reasons (perhaps among other factors, such as coherence).[30] The border between rationality and irrationality may be vague or precise. I do not take a stance on that issue here.

Architectural Sensitivity makes better sense of the spectrum of encapsulation than Lenience or Austerity does. To illustrate this point, I will review how it treats the cases discussed thus far. First, consider states produced by weakly encapsulated systems, such as Lewis's geographical belief fragments. These belief fragments are weakly encapsulated because once the agent simultaneously attends to both fragments,

28. One could also think of Architectural Sensitivity as modulating the strength of Respond to Reasons. The idea of the strength of a rational requirement is vague without further elaboration, so I instead formulate Architectural Sensitivity in terms of the more precise and familiar idea of degree of irrationality.

29. For arguments that perception, and emotion are imperfectly encapsulated, see O'Callaghan (2019) and Majeed (2020). The best candidates for perfectly encapsulated systems are early sensory subsystems such as early vision (Pylyshyn, 1999), though even early vision undergoes certain types of diachronic perceptual learning (e.g., Ahmadi, 2018).

30. For further discussion of degrees of rationality, see Foley (1992), Karlan (2020), and Staffel (2020).

the barrier to information access is quickly and easily eliminated.[31] Architectural Sensitivity says that because the fragments are weakly encapsulated, Lewis's contradictory beliefs are highly irrational. This aligns with the intuitive verdict discussed in §3.

Take next the outputs of emotion, a moderately encapsulated system. Emotion is moderately encapsulated because while we cannot always immediately reason ourselves out of irrational feelings, over time emotions can respond to reasons through techniques such as monitoring, reassurance, redirection of attention, and therapy. According to Architectural Sensitivity, because emotion is a moderately encapsulated system, recalcitrant emotions are moderately irrational. However, emotions that start out as recalcitrant but eventually respond to reasons through these various techniques will ultimately count as rational in virtue of having responded to reasons over time.

The idea that recalcitrant emotions are moderately irrational in virtue of their failure to respond to reasons makes sense of some of the rational differences between beliefs and emotions. In the example of the two employees experiencing imposter phenomenon discussed above, the employee with unresponsive beliefs seems more irrational than the employee with unresponsive emotions. Both violate the requirement to respond to reasons, but the belief is more irrational than the emotion because the belief is produced by a less encapsulated system. Architectural Sensitivity neatly captures the rational difference between belief and emotion that evaded Lenience and Austerity.

Next consider perception, which is a strongly encapsulated system, at least when considered synchronically.[32] No matter how hard we try, we cannot reason ourselves into seeing a known illusion veridically. Additionally, we may not be able to resist automatically forming a belief endorsing that perception. According to Architectural Sensitivity, known illusions and automatic perceptual beliefs are only very weakly irrational because perception is so strongly encapsulated. This verdict fits with the intuitive consensus that known illusions have little impact on our rationality.[33]

However, when considered diachronically, perception is only moderately encapsulated. Through diachronic perceptual learning, perceptual systems can gradually access information stored in other perceptual subsystems, such as in speech perception (McGurk & MacDonald, 1976; Mitchel et al., 2014), rhythm perception, and flavor perception (Connolly, 2019; O'Callaghan, 2020). Perceptual systems can also gradually access information stored in cognition, such as in chess perception (Chase & Simon, 1973; Leone et al., 2014) and mathematical perception (Landy & Goldstone, 2007; Kellman, Massey, & Son, 2009). Perceptual learning requires repeated experience over time, but it is not rare. It is a central function of perception (Connolly, 2019; Jenkin, 2023b).

Architectural Sensitivity explains the rationality and irrationality that can result from perceptual learning. Consider the chess master who has undergone perceptual learning and so directly sees a castling on the board. Their automatic perceptual belief endorsing this perception is rational because it meets the requirement to respond to reasons by responding to their reasons in favor of the presence of a castling (the rules of chess). In contrast, the resistant amateur player possesses the same reasons (they also know the rules of chess) and looks at the same board, but after scanning the board sees no castling and so believes there is no castling. Their belief fails to respond to their reasons, even over time. According to Architectural Sensitivity, because perception is diachronically moderately encapsulated, the resistant amateur's perceptual belief is moderately irrational. This analysis fits

31. Not all belief fragments are weakly encapsulated. Some are very difficult to merge and hence strongly encapsulated. For discussion see Bendaña & Mandelbaum (2021).

32. Perceptual systems do sometimes respond immediately to certain types of external information, such as linguistic labels (Lupyan & Ward, 2013) or crossmodal sensory data (Welch & Warren, 1980; O'Callaghan, 2020). Most often, though, a training period is required for external information to influence perception (Goldstone, Leeuw, & Landy, 2015).

33. Independent rational requirements govern when it is appropriate to form all-things-considered beliefs about known illusions. My focus here is only on automatic perceptual beliefs.

with the intuitive judgment of the resistant amateur described in §3.

6. Control and Can

Even if one accepts that Architectural Sensitivity accurately predicts the relative rationality of a wide range of encapsulated failures, one might still wonder whether it gets to the heart of the matter. A system's degree of encapsulation often correlates with the agent's degree of control over the system's outputs. Given the central role of control in traditional debates over normative requirements, it is natural to ask whether Architectural Sensitivity reduces to the view that when a belief fails to respond to reasons, the belief is irrational to the degree to which it is under the agent's control.

Despite appearances, a system's degree of encapsulation does not always track an agent's degree of control over its outputs. Vision is synchronically strongly encapsulated, but we can immediately control our visual perceptions by turning our head or closing our eyes. Emotion is moderately encapsulated, but we can control recalcitrant anxiety through distraction (e.g., watching cat videos). These cases involve control over a system's outputs by changing the system's inputs, rather than by changing the system's response to a fixed set of inputs. Nonetheless, they are cases of control over the outputs of encapsulated systems.

One might then wonder whether degree of encapsulation (and hence Architectural Sensitivity) reduces to a different, more specific form of control: control over the information a system uses to process a fixed set of inputs. Cases of the opposite type—uncontrolled yet unencapsulated process—challenge this candidate reduction. For example, beliefs formed by inference in central cognition can be influenced by stereotypes despite our attempts to control such influence (Hamilton & Sherman, 1994). Moral reasoning can automatically draw on consciously or unconsciously stored moral rules without the agent's control (Mallon & Nichols, 2011). Associations between concepts can be made outside the agent's control, even within unencapsulated associative networks. These examples illustrate that even control over the information that a system accesses while processing a fixed set of inputs does not reduce to encapsulation. Architectural Sensitivity is best formulated as a thesis about informational encapsulation itself, rather than as a thesis about control.

A related worry is that Architectural Sensitivity only gets the verdicts about these cases right because accessibility of information (as affected by degree of encapsulation) determines the degree to which an agent possesses a reason. In the case of imposter phenomenon, one might say that employee A, whose emotions fail to respond to their reasons of their awards and achievements, possesses those reasons to a lesser degree than employee B, whose beliefs fail to respond to the same reasons. This implies that A is less irrational for failing to respond to their weakly possessed reasons than B is for failing to respond to their strongly possessed reasons. However, both A and B meet the sufficient condition on possessing reasons discussed in §2: they can use those reasons to guide their inferences and actions. Both A and B can use their knowledge of their accomplishments to infer that they have won more awards than their colleagues, or to list their accomplishments on their CVs. If anything, A uses their reasons to guide their inferences and actions more than B does, because A uses their reasons to inform beliefs about their self-worth, while B does not. Thus, A does not straightforwardly possess the relevant reasons less than B.

If Architectural Sensitivity does not reduce to degrees of reasons possession or to degrees or control, one might worry that it places undue demands on agents. Like Austerity, Architectural Sensitivity allows that mental states can be irrational for factors beyond our control. For example, recalcitrant emotions are irrational even though they arise due to architectural borders between emotion and cognition. While Architectural Sensitivity mitigates the irrationality of encapsulated failures, it nonetheless allows that we can be required to respond to a reason when we seemingly cannot do so, challenging 'ought' implies 'can'.

Yet as the cases discussed here demonstrate, encapsulated systems very often *can* respond to reasons, albeit in indirect and subtle ways.

With time, experience, and training, even strongly encapsulated perceptual systems can incorporate new information into their proprietary processing database (e.g., Goldstone, Leeuw, & Landy, 2015). Reflecting this duality of encapsulation and malleability is one of Architectural Sensitivity's central aims. While Architectural Sensitivity sets the standard for rationality high, it is achievable by human minds.

The idea that states produced by psychological systems outside our direct, immediate control (but perhaps within indirect or mediate control) can be rationally required to respond to reasons is not unique to Architectural Sensitivity. Consider in-group bias, a deep-rooted involuntary tendency to favor one's in-group (e.g., members of the same race, gender, political party, or country) over one's out-group (Tajfel & Turner, 1986). While this bias may be evolutionarily advantageous, it is also irrational. It generates unfounded prejudice, limits our evidence pools, and creates unjustified conflict. Despite its innate and instinctual nature, in-group biases seem required to respond to our reasons to reduce them.

Similarly, beliefs produced by cognitive dissonance are not under our immediate control yet are required to respond to reasons. Cognitive dissonance is a reasoning pattern in which we repress counterevidence to our core beliefs and subsequently increase our confidence in those core beliefs rather than revising them (Aronson, 1969). Cognitive dissonance is irrational—it leads to an increase in confidence in the exact beliefs for which our reasons support decreasing confidence (Quilty-Dunn, in preparation). We may be able to find and eliminate individual instances of cognitive dissonance in our own minds, but it is a fundamental feature of human thought that cannot be eradicated wholesale. The fact that cognitive dissonance is built-in and difficult to alter does not exempt it from the requirement to respond to reasons and the irrationality it brings with it. The same is true of the outputs of encapsulated systems.

7. Conclusion

I have argued here that Architectural Sensitivity better accounts for the rationality of the spectrum of encapsulated failures than Lenience or Austerity. The rational status of a belief is modulated by the degree of encapsulation of the system that produced it. Yet because our cognitive systems are rarely perfectly encapsulated—they are typically changeable through indirect interventions or diachronic learning—encapsulated failures are almost always irrational to some degree. This may seem surprising, but it is the thread that best connects our intuitions about the individual examples.

My argument for Architectural Sensitivity is an inference to the best explanation rather than a deductive argument. It is possible that Lenience or Austerity is correct, and our intuitions do not track the rational requirements. It is also possible that an unknown fourth view accounts for our intuitions just as well or better. Given the current landscape, though, Architectural Sensitivity is the best option.

Architectural Sensitivity not only delivers plausible verdicts about the spectrum of encapsulation, but it also incorporates the central insights of both Lenience and Austerity. Lenience was motivated by the idea that rational requirements should be sensitive to our human limitations. Architectural Sensitivity respects this idea by indexing degree of irrationality to the degree of encapsulation of a system. Yet in the spirit of Austerity, Architectural Sensitivity allows that beliefs can be required to do things they cannot immediately do. Beliefs can be irrational for failing to respond to reasons when our mental systems have yet not engaged in the right sort of learning processes to enable a rational response, as in the cases of belief fragmentation, recalcitrant emotion, and resistance to perceptual learning. Architectural Sensitivity captures the idea that while informational encapsulation places major constraints on our ability to respond to reasons, those constraints are

fundamental parts of our rational characters.[34]

References

Ahmadi, M., McDevitt, E. A., Silver, M. A., & Mednick, S. C. (2018). Perceptual learning induces changes in early and late visual-evoked potentials. *Vision Research*, 152, 101-109.

Alston, W. (1988). The deontological concept of epistemic justification. *Philosophical Perspectives*, 2, 257-299.

Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. *Advances in Experimental Social Psychology*, 4, 1-34.

Audi, R. (2008). The ethics of belief: Doxastic self-control and intellectual virtue. *Synthese*, 161, 403-418.

Bendaña, J. & Mandelbaum, E. (2021). The fragmentation of belief. In C. Borgoni, D. Kinderman, & A. Onofri, (Eds.), *The Fragmented Mind* (pp. 78-107). Oxford: Oxford University Press.

Berker, S. (2015). Coherentism via Graphs. *Philosophical Issues*, 25(1): 322-352.

Bilalic, M., Langner, R., Erb, M., & Grodd, W. (2010). Mechanisms and neural bases of object and pattern recognition: A study with chess experts. *Journal of Experimental Psychology: General*, 139(4), 728-742.

Bonjour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Borgoni, C. (2021). Rationality in fragmented belief systems. In C. Borgoni, D. Kinderman, & A. Onofri, (Eds.), *The Fragmented Mind* (pp. 137-155). Oxford: Oxford University Press.

Boyle, M. (2009). Active belief. *Canadian Journal of Philosophy*, 39(S1), 119-47.

34. For helpful feedback and discussion, I thank Ned Block, Allan Hazlett, Eric Mandelbaum, Christopher Peacocke, Susanna Siegel, and Jake Quilty-Dunn. I also thank two anonymous referees at this journal, as well as two anonymous referees at another journal for their comments. I also thank audiences at New York University, York University, the University of Missouri St. Louis, Northeastern University, Johns Hopkins University, and Washington University in St. Louis.

Brady, M. S. (2009). The irrationality of recalcitrant emotion. *Philosophical Studies*, 145(3), 413-430.

Brady, M. S. (2013). *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press.

Campitelli, G., Gobet, F., Head, K., Buckley, M., & Parker, A. (2007). Brain localization of memory chunks in chess players. *International Journal of Neuroscience*, 117(12), 1641-1659.

Carman, M. (2018). Emotionally guiding our actions. *Canadian Journal of Philosophy*, 48(1), 43-64.

Carr, J. R. (2021). Why ideal epistemology? *Mind*, 131(524), 1131-1162.

Charness, N., Reingold, E.M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition*, 28(8), 1146-1152.

Chase, W., and Simon, H. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55-81.

Cherniak, C. (1983). Rationality and the structure of human memory. *Synthese*, 57(2), 163-186.

Cherniak, C. (1986). *Minimal Rationality*. Cambridge, MA: MIT Press.

Chignell, A. (2018). The ethics of belief. *The Stanford Encyclopedia of Philosophy* (2018 ed.). E. N. Zalta (Ed.). URL = <https://plato.stanford.edu/archives/spr2018/entries/ethics-belief/>.

Clance, P. R. & Imes, S. A. (1978). The imposter phenomenon in high-achieving women: Dynamics and therapeutic intervention." *Psychotherapy Theory, Research, and Practice*, 15(3), 241-247.

Clifford, W. K. (1877/1999). The ethics of belief. In T. Madigan (Ed.), *The Ethics of Belief and Other Essays* (pp. 70-96). Amherst, MA: Prometheus.

Conee, E. & Feldman, R. (1985). Evidentialism. *Philosophical Studies*, 48(1), 15-34.

Cosmides, L. & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis and J. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed.) (pp. 91-115). New York: The Guilford Press.

D'Arms, J. & Jacobson, D. (2000). The moralistic fallacy: On the 'appro-

priateness' of emotions. *Philosophical and Phenomenological Research*, 61(1), 65-90.

D'Arms, J. & Jacobson, D. (2003). The significance of recalcitrant emotion. *Royal Institute of Philosophy Supplement*, 52, 127-145.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23), 685-700.

Davidson, D. (1982/2004). Paradoxes of irrationality. In D. Davidson, *Problems of Rationality* (pp. 169-188). Oxford: Oxford University Press.

Davidson, D. (1983/2001). A coherence theory of truth and knowledge. In D. Davidson, *Subjective, Intersubjective, Objective* (pp. 137-154). Oxford: Oxford University Press.

de Sousa, R. B. (1987). *The Rationality of Emotions*. Cambridge, MA: MIT Press.

Egan, A. (2008). Seeing and believing: Perception, belief formation, and the divided mind. *Philosophical Studies*, 140(1), 47-63.

Elga, A. & Rayo, A. (2021). Fragmentation and Information Access. In C. Borgoni, D. Kinderman, & A. Onofri, (Eds.), *The Fragmented Mind* (pp. 37-53). Oxford: Oxford University Press.

Elga, A. & Rayo, A. (2022). Fragmentation and logical omniscience. *Noûs*, 56(3), 716-741.

Fantl, J. & McGrath, M. (2009). *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Feldman, R. (2000). Voluntary belief and epistemic evaluation. In M. Steup (Ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue* (pp. 77-92). Oxford: Oxford University Press.

Fodor, J. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (1983). *Modularity of Mind*. Cambridge, MA: MIT Press.

Foley, R. (1992). *Working Without a Net: A Study of Egocentric Epistemology*. New York: Oxford University Press.

Frankfurt, H. (1973). Coercion and moral responsibility. In T. Honderich (Ed.), *Essays on Freedom of Action* (pp. 65-86). London: Routledge and Keegan Paul.

Friedman, J. (2020). The epistemic and the zetetic. *The Philosophical Review*, 129(4), 501-536.

Georgeson, M., May, K., Freeman, T., & Hesse, G. (2007). From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision*, 7(13), 1-21.

Gertler, B. (2021). Rational agency and the struggle to believe what your reasons dictate. In C. Borgoni, D. Kinderman, & A. Onofri, (Eds.), *The Fragmented Mind* (pp. 325-349). Oxford: Oxford University Press.

Gilbert, D. (1991). How mental systems believe. *American Psychologist*, 46(2), 107-119.

Griffiths, T., Kemp, C., & Tenenbaum, J. (2008). Bayesian models of cognition. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 59-100). Cambridge, UK: Cambridge University Press.

Goldberg, S. (2017). Should have known. *Synthese*, 194(8), 2863-2894.

Goldstone, R. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585-612.

Goldstone, R., Leeuw, J. R., & Landy, D. H. (2015). Fitting perception in and to cognition. *Cognition*, 135, 24-29.

Hamilton, D. L. & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, 103(2), 336-355.

Helton, G. (2020). The revisability view of belief. *Noûs*, 54(3), 501-526.

Hursthouse, R. (1991). Arational actions. *The Journal of Philosophy*, 88(2), 57-68.

James, W. (1896/1956). *The Will to Believe and Other Essays in Popular Philosophy*. New York: Dover Publications.

Jenkin, Z. (2023a). Perceptual learning and reasons-responsiveness. *Noûs*, 57(2), 481-508.

Jenkin, Z. (2023b). The function of perceptual leaning. *Philosophical Perspectives*, 37(1), 172-186.

Jones, K. (2003). Emotion, weakness of will, and the normative conception of agency. *Royal Institute of Philosophy Supplement*, 52, 181-200.

Karlan, B. (2020). Reasoning with heuristics. *Ratio*, 34(2), 100-108.

Kant, I. (1787/1997). *Critique of Practical Reason*. M. Gregor (Ed. and Trans.). Cambridge, UK: Cambridge University Press.

Kellman, P. J., Massey, C. M., & Son, J. (2009). Perceptual learning modules in mathematics: enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, 2(2), 285–305.

Kelly, T. (2016). Evidence. In *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed), E. N. Zalta (Ed.). URL = <https://plato.stanford.edu/archives/win2016/entries/evidence/>.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 720–733.

Leone, M. J., Slezak, D. F., Cecchi, G. A., & Sigman, M. (2014). The geometry of expertise. *Frontiers in Psychology*, 5(47), 1-9.

Lehrer, K. (2000). *Theory of Knowledge* (2nd ed). Boulder: Westview Press.

Lewis, D. (1982). Logic for equivocators. *Noûs*, 16(3), 431-441.

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology*, 6(4), 547-569.

Lupyan, G., & Ward, E. (2013). Language can boost otherwise unseen objects to visual awareness. *Proceedings of the National Academy of Sciences*, 110(35), 14196-14201.

Lyons, J. (2011). Circularity, reliability, and the cognitive penetrability of perception. *Philosophical Issues*, 21(1), 289-311.

Macpherson, F. (2009). Perception, philosophical perspectives. In T. Bayne, A. Cleeremans, & P. Wilken (Eds.), *The Oxford Companion to Consciousness* (pp. 502-508). Oxford: Oxford University Press.

Maguire, B. (2018). There are no reasons for affective attitudes. *Mind*, 127(507), 1779-805.

Majeed, R. (2019). What can information encapsulation tell us about emotional rationality? In L. Candiotto (Ed.), *The Value of Emotions for Knowledge* (pp. 502-508). Basingstoke: Palgrave Macmillan.

Majeed, R. (2020). Does modularity undermine the pro-emotion consensus? *Mind and Language*, 35(3), 277-292.

Mallon, R. & Nichols, S. (2011). Dual processes and moral rules. *Emotion Review*, 3(3), 284-285.

Mandelbaum, E. (2013). Thinking is believing. *Inquiry: An Interdisciplinary Journal of Philosophy*, 57(1), 55-96.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.

Moran, R. (1988). Making up your mind: Self-interpretation and self-constitution. *Ratio*, 1(2), 135-151.

Moran, R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.

Mitchel, A. D., Christiansen, M. H., Weiss, D. J. (2014). Multimodal integration in statistical learning: Evidence from the McGurk illusion. *Frontiers in Psychology*, 5(407), 1-6.

Naar, H. (2021). Skepticism about reasons for emotions. *Philosophical Explorations*, 25(1), 108-123.

O'Callaghan, C. (2019). *A Multisensory Philosophy of Perception*. Oxford: Oxford University Press.

Pasnau, R. (2013). Epistemology idealized. *Mind*, 122(488), 987-1021.

Pavese, C. (2025). The epistemology of skill. In K. Sylvan (Ed.), *Blackwell Companion to Epistemology* (3rd ed.) (chapter 43). Chichester: John Wiley & Sons Ltd.

Prinz, J. (2004). *Gut Reactions: A Perceptual Theory of Emotions*. Oxford: Oxford University Press.

Prinz, J. (2006). Is the mind really modular? In R. Stainton (Ed.), *Contemporary Debates in Cognitive Science* (pp. 22-36). Malden: Blackwell Publishing.

Pylyshyn, Z. W. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge: MA, MIT Press, A Bradford Book.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22(3), 341-423.

Quilty-Dunn, J. (2015). Believing in perceiving: Known illusions and the classical dual-component theory. *Pacific Philosophical Quarterly*,

96(4), 550-575.
Quilty-Dunn, J. (2022). Unconscious rationalization, or how (not) to think about awfulness and death. [Manuscript in preparation]. https://philpapers.org/archive/QUIURO.pdf
Quilty-Dunn, J. & Mandelbaum, E. (2018). Against dispositionalism: Belief in cognitive science. *Philosophical Studies*, 175(9), 2353-2372.
Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review*, 60(1), 20-43.
Rinard, S. (2018). Believing for practical reasons. *Noûs*, 53(4), 763-784.
Saariluoma, P. (1991). Aspects of skilled imagery in blindfolded chess. *Acta Psychologia*, 77(1), 65-89.
Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford: Oxford University Press.
Scanlon, T. M. (2015). Forms and conditions of responsibility. In R. Clarke, M. McKenna, and A. Smith (Eds.), *The Nature of Moral Responsibility: New Essays* (pp. 89-105). Oxford: Oxford University Press.
Schultz, C. (2025). Deliberative control and eliminativism about reasons for emotions. *Australasian Journal of Philosophy*, 103(1), 72-87.
Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Noûs*, 36(2), 249-275.
Shah, N. (2006). A new argument for evidentialism. *The Philosophical Quarterly*, 56(225), 481-498.
Siegel, S. (2017). *The Rationality of Perception*. Oxford: Oxford University Press.
Slank, S. (2019). Rethinking the imposter phenomenon. *Ethical Theory and Moral Practice*, 22(1), 205-218.
Smith, H. (2014). The subjective moral duty to inform oneself before acting. *Ethics*, 125(1), 11-38.
Smortchkova, J. (2017). Encapsulated social perception of emotional expressions. *Consciousness and Cognition*, 47, 38-47.
Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Malden, MA: Blackwell.
Staffel, J. (2017). Should I pretend I'm perfect? *Res Philosophica*, 94(2), 301-324.
Staffel, J. (2020). *Unsettled Thoughts*. Oxford: Oxford University Press.
Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.
Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.). *Psychology of Intergroup Relations* (pp. 7-24). Chicago: Nelson Hall.
Van Inwagen, P. (1996). It is wrong, everywhere, always, and for anyone, to believe anything upon insufficient evidence. In J. Jordan & D. Howard-Snyder (Eds.). *Faith, Freedom, and Rationality* (pp. 137-153). Lanham, MD: Rowman and Littlefield.
Vranas, P. (2007). I ought, therefore I can. *Philosophical Studies*, 136(2), 167-216.
Watzl, S. (2017). *Structuring Mind: The Nature of Attention and How it Shapes Consciousness*. Oxford: Oxford University Press.
Weatherson, B. (2008). Deontology and Descartes' demon. *Journal of Philosophy*, 105(9), 540–569.
Wedgewood, R. (2013). Rational 'ought' implies 'can.' *Philosophical Issues*, 23(1), 70-92.
Welch, R. B. & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638-667.
Williamson, T. (2000). *Knowledge and its Limits.* Oxford: Oxford University Press.
Worsnip, A. (2018). The conflict of evidence and coherence. *Philosophical and Phenomenological Research*, 96(1), 3-44.
Yalcin, S. (2021). Fragmented but rational. In C. Borgoni, D. Kinderman, & A. Onofri, (Eds.), *The Fragmented Mind* (pp. 156-179). Oxford: Oxford University Press.