



# Longings in Limbo: A New Defence of I-Desires

Luke Roelofs<sup>1</sup>

Received: 6 May 2021 / Accepted: 3 December 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

This paper responds to two arguments that have been offered against the positing of ‘i-desires’, imaginative counterparts of desire supposedly involved in fiction, pretence, and mindreading. The Introspection Argument asks why, if there are both i-desires and desires, the distinction is so unfamiliar and hard to draw, unlike the relatively clear distinctions between perception and mental imagery, or belief and belief-like imagining. The Accountability Argument asks how it can make sense to treat merely imaginative states as revealing of someone’s psychology, the way we do with responses to fiction. I argue that carefully considering the relationship between other states and their imaginative counterparts sheds light on how we should expect i-desires to differ from desires, and suggests that we may often be in states that are *indeterminate*, in limbo between the two categories. This indeterminacy explains why the distinction is often hard to draw, and why these states can be revealing about us even without (determinately) being real desires.

Some things we believe to be true, while others we only imagine to be true. Some things we see with our eyes, while others we see only with ‘the mind’s eye’. Some theorists have suggested an extension of the pattern: some things we really desire, while others we only ‘i-desire’.<sup>1</sup> An i-desire is a desire-like imagining: a species of imagining alongside visual imagining, or belief-like imagining (aka ‘i-belief’), or any other sort (auditory, motor, olfactory, etc.). Just as an image in my mind’s eye is very like a visual perception but differs in important ways, and a belief-like imagining is very like a belief but differs in important ways, so an i-desire is very like a desire but differs in important ways.

<sup>1</sup> For defences of something like this idea, see Currie 1997, 2002, Walton 1997, Velleman 2000, Currie and Ravenscroft 2002, Doggett and Egan 2007. The term ‘i-desire’ is from Doggett and Egan; other authors speak of ‘desires’<sup>1</sup>, ‘desire-like imaginings’, ‘offline desires’, ‘wishes’, or use other terms, but the essential idea of a conative state that is an imaginative counterpart to regular conative states is shared.

✉ Luke Roelofs  
luke.mf.roelofs@gmail.com

<sup>1</sup> New York University, New York City, USA

To use a popular example in the literature (originating with Doggett & Egan, 2007, p.2 ff), we might consider viewers of *The Sopranos*, who perhaps in reality desire that anyone like Tony Soprano be apprehended, held accountable, and stripped of their power to harm others, but who nevertheless, while watching Tony Soprano trying to escape arrest, feel anxiety on his behalf, and hope that he gets away. Rather than saying that they really desire that he escape, philosophers have suggested, we should say that these viewers i-desire that he escape: they are in a state that feels and behaves a lot like desire, but that is disengaged from their real commitments in the same way as the other imaginings (visual, belief-like, etc.) that let them engage with the show while knowing it's not real.

The postulation of i-desires has been met with a range of objections,<sup>2</sup> to which defenders have offered various replies.<sup>3</sup> In this paper I aim to add to those replies, and in particular to defend i-desires against a particular line of objection that as far as I know has not yet been directly addressed. Doggett and Egan, in defending i-desires against a host of other objections, admit the following:

[O]rdinary people... don't call how they feel about Tony Soprano "i-wanting" him to be well. If anything, they'd call it "wanting." Why is there no folk psychological notion of... i-desire? We have a folk psychological notion of the imagination. Why not other imaginative analogues? These are interesting questions to which we have no answer. (Doggett & Egan, 2012, p.304)

The point is not just about what words happen to be in our language; it is about the concepts and distinctions that we employ in self-understanding. Kind, a critic of i-desires, elaborates on the same theme:

Typically, we can tell whether we are believing something or merely imagining it – we don't mistake our belief-like imaginings for beliefs. But we don't seem to have any way to tell whether we are desiring something or merely i-desiring it... since we normally take ourselves to be having genuine desires while, say, engaging with fiction, the postulation of desire-like imagination requires that we are systematically mistaken about our own mental states." (Kind 2016a, p.173, cf. Kind, 2011, pp.30-31)

If there are i-desires, just as distinct from real desires as belief-like imaginings are from beliefs, why is this not a familiar, common-sense point? And why is it not obvious which one we're having on a given occasion? Call this the Introspection Argument.

I argue that the different functional roles of beliefs and desires naturally imply much more scope for desire-like states to blur the boundary between real state and imaginative counterpart. The possibility of states that blur this boundary—that are half-real, half-imaginative, or as I will say 'indeterminately offline'—has already been recognised and argued for (see, e.g., Egan, 2009, Schellenberg, 2013, cf.

<sup>2</sup> See, e.g., Stich and Nichols (2003), Weinberg and Meskin (2005), Funkhouser and Spaulding (2009), Kind (2011, 2016a), Tagliafico (2011), Langland-Hassan (2020).

<sup>3</sup> See, e.g., Currie (2010), Doggett and Egan (2011).

Schwitzgebel, 2001) in the case of beliefs. Sometimes we seem to half accept something and half not, such that it is unclear whether we are just imagining it or really believing it. But with belief these cases are both fairly rare and presumptively irrational—whether we’re treating a belief like an imagining, or treating an imagining like a belief, we seem to be falling short of perfect epistemic rationality.

I think indeterminately offline desires, by contrast, are unnoticed in part because of their frequency: they may even outnumber determinately offline desires. Moreover, there is nothing irrational about them. As long as we act and reason rightly, there is nothing wrong with letting our desire-like states linger in a limbo between being real desires and being i-desires.

This also helps defend i-desires against another objection raised by Kind, concerning the revealingness of desires in our response to fiction:

We typically hold one another accountable for the kinds of responses we have to fiction... Given that I am not generally sympathetic to mafia members, I thus might be puzzled as to why I want Tony Soprano to escape capture by the police... such evaluative practices make little sense unless the conative states we are evaluating are instances of genuine desire. (Kind, 2016a, p.173, cf. Kind, 2011, pp. 30–31)

Call this the Accountability Argument. It makes an important point, though it can be overstated. Sometimes people are inclined to morally judge, or at least look critically at, desire-like states experienced in response to fiction, and even more so those that govern private fantasy; but sometimes people are inclined to reject such judgement, to insist that what we want ‘in our imagination’ is a different matter from what we ‘really want’, and should be judged by different standards if judged at all.<sup>4</sup> If it is fair to worry that i-desire theorists cannot make sense of the first fact, is also fair to worry that their opponents cannot make sense of the second. My hope is that explaining why the boundary between real desires and i-desires can get so blurry, we can also better explain our tendencies both to judge and to withhold judgement.

In essence my account is this: imagination allows us to lift the restrictions that usually regulate our mental life—to ‘see’ what is not before us, to ‘believe’ things inconsistent with our evidence, and to ‘desire’ things we recognise as undesirable. But desire is by its nature less tightly restricted than either belief or perception, not constrained by environmental stimuli or by rules of evidence, and so the lifting of restrictions has a less pronounced effect. Instead of two sharp categories, the real and the imaginative, the result is a continuum of desire-like states whose status, and eligibility for moral judgement, are often genuinely indeterminate, and even when determinate may be hard to ascertain.

<sup>4</sup> For discussion see, e.g., Hazlett (2009), Stear (2009), Smuts (2015, 2016).

## 1 The Puzzle of Imaginative Desire

The anchoring points of this debate, as it has developed in the literature, are fiction and pretence. It seems that our reactions to both involve something desire-like: not only do we naturally talk about what we want to happen in a story or in a game we are playing, but we feel affective reactions to fictional events that seem to reveal preferences about them, and perform pretend-actions that seem to be motivated by some sort of pretend-goal.<sup>5</sup>

Consider a few examples:

- Watching *The Sopranos*, I might ‘want’ Tony Soprano to escape the police. (Doggett & Egan, 2007, p.12)
- Watching the movie *JFK*, I might ‘want’ the prosecution of supposed conspirator Clay Shaw to succeed, even though based on my factual beliefs that it was in fact baseless, I am glad it failed (Friend, 2003, p.41 ff).
- Watching *Jurassic Park*, I might ‘want’ the raptors to survive the final fight scene with the T. rex, eat the humans, and live happily on the island.
- Reading *Les Misérables*, I might ‘want’ the June Rebellion of 1832 to succeed.
- Playing a game where I pretend to be a terrible biting monster, I might ‘want’ to bite my loved ones, but instead just mime out a biting action (Kind, 2011, p. 436).

Three major analyses of these conative states have been staked out (Currie, 2010, cf. Langland-Hassan 2020, pp. 212–214). First, the ‘simple view’ says that they are simply desires, and that their content is the sort of first-order fictional statements that we would naturally use to express them (see Kind, 2011; 2016a, Spaulding, 2015). If someone says ‘I hope Tony Soprano gets away’, they are expressing a desire whose content is simply ‘Tony Soprano gets away’.

Second, the complex content view agrees that what we have here are genuine desires, but says that their content contains an implicit reference to the fiction itself (see Nichols, 2004, p.332, Langland-Hassan 2019). The desire that Tony get away is really a desire that this TV show, *The Sopranos*, be such that in it, Tony gets away. It is not a desire regarding a New Jersey mob boss (after all, that mob boss doesn’t exist), but a desire regarding a TV show.

Third, the i-desire view agrees with the simple view that our desire-like states are directed at fictional things, not at the fiction itself, but claims that at least sometimes, these states are not genuine desires, but i-desires. We don’t really *desire* Tony

<sup>5</sup> Many in this debate, following especially Walton (1990), treat pretence and fiction as continuous with one another: two ways to create a shared imaginary world, differing simply in what sort of actions are taken by the participants. Reading a novel, on this analysis, is engaging in a certain sort of pretence: pretending that the author is telling a true story, and pretending to believe it. Although some dispute this analysis (e.g. Langland-Hassan, 2020, p.144 ff), I find it congenial, particularly because it allows for a unified discussion of what is ‘true in’ a fiction or a game. So I will assume for convenience that references to ‘fictional truth’ or ‘fictional characters’ also cover truth in a game and characters in a game. Those who reject this analysis can mentally substitute ‘fictional-or-pretend’ as necessary.

Soprano to get away: rather we i-desire it, i.e. create an imaginative counterpart of desiring it. The analogy is to how we do not really *believe* that Tony Soprano lives in New Jersey: we imagine it, aka i-believe it, create an imaginative counterpart of believing it.<sup>6</sup>

Adjacent to this debate about *desire* is the long-running debate about the *emotions* we feel in response to fiction; roughly corresponding positions can be sketched out there, that we either feel genuine emotions directed at fictional things, genuine emotions directed at the fiction as such, or imaginative analogues of real emotion (sometimes called ‘quasi-emotions’).<sup>7</sup> Since emotions and desires are so intertwined, many considerations bearing on one debate translate to the other, and the two are sometimes not sharply distinguished (e.g. Walton, 1990, pp.258–59, Friend, 2003, pp.49–51). Indeed, proponents of i-desires often appeal to them precisely to explain emotional reactions to fiction (e.g. Currie, 2010; Doggett & Egan, 2012).<sup>8</sup> But the two issues need not always line up (Currie and Ravenscroft, for instance, posit imaginative counterparts of desires but not emotions 2002 pp.189–191), and as we will see in the next section, the two phenomena sometimes come apart in interesting ways.<sup>9</sup>

Many arguments for and against these different views have been offered, and I will not here attempt to survey them all, since my focus is on addressing the Introspection and Accountability arguments. But sketching some of the lines of argument can clarify the shape of the three competing views, so I will try to do so concisely.

First, proponents of the simple view and the i-desire view can argue, against the complex content view, that although we do often have the kind of desires it posits, directed at works of fiction as such, they are recognisably different from the conative states in the above examples. We often want a TV show, a novel, or a game, to be longer, shorter, different in focus, or just better in various ways. But there is a phenomenological difference between thinking about the fiction *as* a fiction, wishing it were more serious or more light-hearted or whatever, and thinking about the

<sup>6</sup> It would be fairly natural to say that what we i-desire, we ‘imagine desiring’, and in a sense this is true: we are engaged in a certain sort of imagination which resembles desiring. But the phrase ‘imagine desiring X’ can be misleading, since it could be read as ‘imagine *that* we desire X’—i.e. i-believe that we desire X, imaginatively recreate the *belief* that we desire X. This is different from i-desiring X, just as believing that we desire something is different from desiring it.

<sup>7</sup> For discussion of quasi-emotions see esp. Walton (1978, 1990, 1997), Friend (2003, 2020); for opposing views see e.g. Radford (1975), Carroll (1990), Gaut (2003), cf. Dorsch (2011).

<sup>8</sup> Partly for this reason, I will in this paper use ‘desire’ broadly to cover attitudes we may have towards past events, which we cannot change or even hope for a different outcome to. If I desire that the June Rebellion succeed, that desire cannot be satisfied, but I still feel it. Some authors might restrict the term ‘desire’, and say what I have in this case is really a ‘wish’, or an emotion of regret or sadness: this will not substantially affect any of my arguments.

<sup>9</sup> Another adjacent debate is about the causes and nature of imaginative *resistance*, our tendency to simply refuse to ‘go along with’ certain propositions we are asked to imagine, often seemingly because we find them repugnant in a way that merely false, or even self-contradictory, propositions are not (see, e.g., Gendler 2000; Liao et al., 2014). I think i-desires may play a role in this phenomenon, but I am doubtful that their role is essential: I am persuaded by Gendler’s argument (2000, pp.73–75) that purely factual propositions may elicit imaginative resistance when their connection to hateful beliefs in the real world is sufficiently salient.

characters and events within the fiction, e.g. wishing Tony were less of a psychopath. This difference is often gestured at with talk of ‘immersion’: when we think about the fiction as such, we have a sort of distance from it, but when we get immersed in it, we seem to think about the fictional things themselves.

Moreover, these two sorts of conative states can conflict in characteristic ways, particularly in tragedy, horror, and other ‘unpleasant’ fictions. Viewers of *Othello* or *Romeo and Juliet* may have a settled, unambiguous desire for tragic fictions. But it’s inherent in a fiction being tragic that it makes us sad, by showing the destruction of things and people we have come to care about, i.e. have come to want not to be destroyed. So when the tragic hero dies, it has to be frustrating our immersed, inside-the-fiction desire-like states: but this does not mean we want the fiction, considered as a fiction, to be different. Even outside tragedy, we often want a story to be ‘challenging’, in not immediately giving us/the protagonists everything we want all the time—pure wish-fulfilment fantasies have their place, but often that’s not what we’re after. So it seems we must distinguish fiction-directed desires from immersed conative states: the simple view and the i-desire view can readily do so, while the complex content view struggles to.

On the other hand, proponents of both the i-desire view and the complex content view can argue, against the simple view, that although we may often have the kind of genuine desires it posits, their functional profile makes them recognisably different from the conative states in the above examples. Desires are defined, in part, by driving us towards action, and yet desires in fiction do not—we have no tendency at all to try to help Soprano escape. Even in games of pretend where we do act out our ‘desires’, we do not do so as we would if we *really* desired them: the terrible biting monster ‘desires’ to bite, not to mime biting, and yet the child only mimes biting. Moreover, desires are often sensitive to practical or moral commitments, and usually abandoned if we come to believe that their objects do not exist. Yet i-desires seem mostly untouched by such factors—not always (as the Accountability Argument brings out) but often. In particular, we would normally stop desiring that someone succeed if we learnt that there was no such person, yet the same knowledge seems to do nothing to our conative reactions to fiction.

Of course, sometimes those reactions are influenced by our real desires, and sometimes a state that first arises in connection with fiction comes to influence our real-world attitudes and behaviour. But to properly analyse what is going on in these cases, we need to be able to distinguish between a real desire and a state whose import is restricted to a movie, game, or novel. The complex content view and the i-desire view can draw this distinction, while the simple view struggles to.

(I-desire theorists sometimes claim that *all* conative states that have fictional objects or events in their content must be i-desires, simply because their objects are fictional. I am unsure that this strong claim is supported by the arguments given for it, but even if it is, the distinction I draw here does not go away: it simply becomes the distinction between i-desires that arise from directly ‘importing’ our real desires into an imaginative project, and those that don’t. This distinction is clear with belief-like imagining: in a sense we ‘imagine that’ humans have kidneys when we read Sherlock Holmes stories, just because we believe that and incorporate it into our imaginative project. We could say that we thereby create a new imaginative state, an

i-belief with the content ‘humans have kidneys’, or we could deny this and say that we’re just bringing our real belief ‘into’ the imagining. Whichever form of description we prefer, the distinction between *this* sort of i-belief, directly derived from real belief, and i-beliefs that go beyond or contradict real beliefs, is still a clear and important distinction.

I think the conjunction of the above two arguments supports the i-desire view: the first supports it against the complex content view, while the second supports it against the simple view. But it enjoys these two advantages only because it posits a greater range of states: fiction-directed desires, real desires directed at fictional objects, and i-desires. The other two views posit fewer (the complex content view only one, the simple view only two), and so might suggest that the i-desire view is unparsimonious, needlessly multiplying mental state types (cf. Funkhouser & Spaulding, 2009, p.299, Kind, 2011, p.423, p.429, Langland-Hassan, 2020, p.214 ff).

While this argument has force against some proponents of i-desires, I think there is an easy response available for those who accept ‘recreativism’, the idea that the imagination is a capacity to produce ‘recreations’ of many types of state, such that the diversity of other mental states will be partly or fully replicated in a diversity of imaginative states.<sup>10</sup> Currie and Ravenscroft put things this way:

Imaginative projection involves the capacity to have, and in good measure to control the having of, states that are not perceptions or beliefs or decision or experiences of movements of one’s body, but which are in various ways like those states—like them in ways that enable the states possessed through imagination to mimic and, relative to certain purposes, to substitute for perceptions, beliefs, decisions, and experiences of movements. (Currie & Ravenscroft, 2002, p.11, cf. Goldman 2006a, 2006b, p.46)

This is sometimes put by calling imaginings ‘offline copies’ of other mental states: visualising a tiger is running our visual system offline to form an offline copy of seeing a tiger; imagining that tigers are secretly robots is running our belief system offline to form an offline copy of believing that tigers are secretly robots, and imagining running from a tiger-robot is running our motor system offline to form an offline copy of the motor program we would use to actually run away from a tiger-robot. The states which are being recreated here can by contrast be called the ‘online counterparts’ to the imaginings.<sup>11</sup>

<sup>10</sup> This view has been called both ‘recreativism’ (Balcerak Jackson, 2018, p.216 ff, Tagliafico 2011, pp.63–76, cf. Currie and Ravenscroft 2002, p.11 ff,) and ‘simulationism’ (e.g. Goldman 2006a, 2006b; Kind 2013, p.4 ff, Balcerak Jackson 2018, p.216 ff). I prefer the former term to avoid confusion with simulationism about mindreading.

<sup>11</sup> This way of distinguishing ‘online’ from ‘offline’, as roughly capturing ‘real’ vs. ‘imaginative copy’, is importantly different from another way of using the terms, on which ‘offline’ cognition is a broader category covering everything not focused on our current environment and activities, including quite genuine beliefs about, e.g. mathematics or the future. Lu Teng has my gratitude for pushing me to clarify this point. Moreover, ‘offline’ here is not intended to imply a specific picture of the neural implementation of the states involved in imagination (as in, e.g. Heal 1998); it is simply an adjective that means the same as the prefix ‘i-’ in i-desires, i-beliefs, etc.

Recreativists can respond thus to the charge of unparsimoniousness: *i*-desires are not a new posit, but simply a natural consequence of the imagination's general capacity to recreate other states. Given the existence of *i*-beliefs as well as beliefs, and *i*-perceptions (i.e. mental imagery) as well as perceptions, allowing for *i*-desires as well as desires is just extending the pattern. Indeed, it would be unparsimonious to *reject* *i*-desires, since we would need a special explanation of why the pattern breaks down here.

There are other arguments that have been made on each side of the debate over desire in imagination, which I cannot do justice to here. In particular, there are challenges for any view in consistently assigning definite contents to the relevant states, whether we think they are genuine desires or *i*-desires (see, e.g., Nichols & Stich, 2003, pp.37–38, Doggett & Egan, 2007, pp.5–8, 13–15, Kind, 2011, pp.425–426, 428–429, 436–437). But the above, I hope, serves to convey some of the major contours of the three approaches, and the basic case for *i*-desires: we have many desire-like states that seem 'immersive' in a way that contrasts with fiction-directed desires, but also seem 'offline' in a way that contrasts with real desires, and if recreativism is accepted as a general account of the imagination, *i*-desires are a natural extension of it.

But this just underlines the force of the Introspection Argument. All the above seems to apply in exactly the same way to, for instance, belief-like imaginings: they are neither beliefs about a fiction nor real beliefs about fictional objects. So why is there such a big difference when it comes to our introspective access to the real/recreation distinction? Why are *i*-beliefs so easy to separate from beliefs, but *i*-desires so hard to separate from desires?

## 2 *I*-Desires Outside of Fiction and Pretence

This debate has been focused primarily on our engagement with fiction and pretence—that is, with imagination in its 'transcendent' use (Kind & Kung, 2016, p.1 ff), imagining things beyond our own reality, that we know to be unreal and engage with nevertheless. But it also touches importantly on how we should think about imagination in what Kind and Kung call its 'instructive' use—as a tool to help us understand reality as it is or might be. We seem to use imagination to consider alternative possible futures, to judge what would have happened if things had gone differently, to help us make decisions about what to do, and to understand others' perspectives.<sup>12</sup> All of these domains are likely to involve desire in one way or another: we may want to judge which plan of action best fits our desires, or think about what we would have done in counterfactual circumstances, or to understand what someone else wants and use it to predict their next move. Whether we should posit *i*-desires to

<sup>12</sup> For discussion of various instructive uses of imagination see, e.g. Yablo (1993), Gordon (1995), Chalmers (2002), Goldman (2006a, 2006b), Kung (2010), Dorsch (2012, 2016), Kind (2016b, 2018), Nanay (2016b), Williamson (2016), Steuber (2016), Van Leeuwen (2016a, 2016b), Balcerak Jackson (2018, Forthcoming), Arcangeli (2018), Egeland (2019), Myers (2021a, 2021b).



explain our engagement with fiction will partly depend on what the best approach is to understanding the role of desire in instructive uses of imagination.

On the one hand, considering instructive uses of imagination can greatly strengthen the case for i-desires, because it seems clear that many of the desire-like states involved cannot be our own genuine desires (cf. Currie, 1995, p.66 ff; Goldman 2006a, 2006b). For example, suppose I am playing chess, and to predict my opponent's next move, I try to imagine the game as they see it, to shift myself into their perspective. A key element of their perspective is their desire to win—that is, their desire that I lose. I have to recreate this desire in order to properly inhabit their perspective, yet I clearly do not really desire that I lose. The i-desire I use to represent their real desire is directly opposed to my actual desire, namely that I win.<sup>13</sup>

There are also many cases where definitely real desires seem to be in play. When I imagine two possible futures to help me decide what to do now, my evaluation of those futures clearly involves real desires. And when I imagine someone else's perspective, I may well find that it motivates me to help them get what they want, i.e. gives rise to real, altruistic, desires. But this is no problem for an i-desire theorist who accepts that both real desires and i-desires can play a role in imagining.

But at the same time, there is potential trouble for the i-desire theorist in the following observation: the cases where it is clearest that our desire-like states are not our real desires are also ones where they produce little or no affect: when imagining my chess opponent's perspective, I may imagine being pleased when I see a crucial opening for them on the board, but I'm not *actually* pleased—the affect I actually feel is dismay, grounded in my real desires which I now see may well be frustrated. In general, it seems I often simulate other people's desires quite 'bloodlessly', feeling nothing on their behalf. But part of the motivation for introducing i-desires in the first place was to explain our affective response to fiction, so doesn't positing them here, not generating any such affect, undermine that initial rationale?

This discrepancy does undermine any simplistic picture of the relationship between affect and conative states, but we should already be sceptical of such a picture. Sometimes perfectly real desires, even ones which I am in the process of acting on, leave me cold. Perhaps I want more than anything to finish my PhD, and am foregoing a raucous board game night in order to do so. But hearing my friends having fun in the next room, the affect I feel is all negative, a product of my desire to join them, not my desire to finish the PhD. The latter desire is perfectly real and genuine; it's even on my mind (I keep reminding myself of it!), but right now it is not determining how I feel. And even when a desire does have an affective pull, the intensity of the affect produced depends a lot on our mindset, our attention, what else is on our mind. A surgeon before an operation might have a problem with their acute awareness of their patient's body as valuable and vulnerable, feelings which make it

<sup>13</sup> It might be objected that here, 'imagining desiring X' is plausibly taken just as meaning 'imagining that I desire X'. But it is not: we can imagine that we have all sorts of properties, mental or physical or relational, but with mental properties there is a clear difference between this and imaginatively recreating the mental state itself—imagining seeing green, unlike simply imagining that I see green, has a visual phenomenology.

harder for them to cut into it. To control this, they have to ‘put aside’ those feelings, so as to see the patient just as a physiological structure—but there is no change here in what they desire (they want the patient to be a happy, healthy, safe, human being, as they did before), just what they are attending to, and how they are attending to it. This effort to ‘distance themselves’ from the humanity of their patient is roughly the same thing we do when we try to push the dangers of a frightening task aside to help ourselves do it, or distance ourselves from our emotional bond with someone when trying to judge impartially between them and someone else.

So we should already accept that other factors can moderate the affective impact of a conative state, so that the same conative state may produce intense feelings or none at all. A good fiction is one that not only makes us *i*-desire things strongly, but also directs our attention, mindset, and receptivity so as to let those desires generate strong emotions. A good thought experiment in a philosophy paper, by contrast, usually aims to do the opposite: to put us into a frame of mind where our reaction to an imagined scenario is largely disinterested cognition without any affect (cf. Peterson, 2021). How exactly this is accomplished is an interesting question, but beyond the scope of this paper; the point is just that the relationship between desires and affect is clearly moderated by various factors both in imagination and outside it.

Indeed, a recreativist might say that we should expect *i*-desires to play different functional roles in different cases, since that is true of imagining generally. Desires, like any other state, can have a range of effects, like focusing attention, motivating behaviour, or evoking affect; when we recreate them in imagination we may be seeking to recreate some or all of these effects, and may succeed more or less fully. The whole point of imaginative recreations, according to recreativism, is that they let us get some of the characteristic effects of a given mental state without others. But this easy recreativist line invites certain difficult questions. If *i*-desires sometimes produce one effect of desire and sometimes others, we might expect recreativism to provide some general guidance about which ones and when. When Currie and Ravenscroft say that recreations are not the same sort of state as their online counterparts, but are “in various ways like those states”, what exactly are the various ways? Ideally, recreativists would give us some sort of rule for how an *X*-like imagining will resemble an *X*, and how it will differ. That would help us evaluate the postulate of *i*-desires, by telling us how to expect *i*-desires to resemble, and differ from, desires. Unfortunately, existing versions of recreativism generally fall short of this kind of generality. While I cannot here fully supply this defect, I think there are some useful observations that can be made that go some way to remedying it.

### 3 What’s Offline?

What, in general, is the difference between an imaginative ‘offline’ version of a state, and its non-imaginative ‘online’ counterpart? And what, conversely, do they have in common? The last two sections both circled this question, in that being able to treat offline states in a unified way was important both for making *i*-desires seem less like an unmotivated new posit, and for understanding the different effects they can have in different circumstances.

Existing discussions often provide key parts of an account of offline-ness, either in general or for specific state-types, without providing a complete general principle. But I think that careful consideration of these partial answers is suggestive of something more general. For example, one obvious feature of imagining is that it is often under voluntary control, in a way that online states aren't. But I think this a symptom of something else, not a defining feature (though see Dorsch, 2012 for an account of imagining that centres voluntariness): though imaginings are often voluntary, they are also often involuntary, both in the weak sense that, once embarked on an imaginative project, it may often evolve in ways that surprise or educate us, and in the strong sense that imaginings may come to us unbidden, influence us subconsciously, and even remain despite our repeated efforts to banish them. (Indeed, i-desires often seem to be among the least voluntary sorts of imagination, as reflected in their frequent invocation to explain the phenomenon of imaginative resistance, cf. Currie & Ravenscroft (2002), Stokes (2006)). Given this, I prefer to take the key thing to be that imaginings are *potentially* open to greater voluntary control than other states, as an aspect of their generally being *less tightly constrained* than other states. But in what sense are they less constrained?

Here is a common gloss: offline states are “disconnected from action-generating systems” (Currie & Ravenscroft (2002), p.66; cf Currie (1997), p.66: offline states “are like real beliefs and desires in terms of internal causal role, but unlike them in terms of external causal role.”) It's true that when I see a tiger or decide that I need to go to the shops, I take appropriate actions (standing up, leaving the room, etc.), but when I merely visualise a tiger or imagine that I need to go to the shops, I don't. But this criterion is not enough by itself. For one thing, imagining can lead to action, particularly in games of pretence, but also when we use imaginative simulation to make better decisions about what to do.<sup>14</sup> For another, imaginings differ from other states in their internal roles too—most obviously, they do not support beliefs the way that real beliefs and perceptions do. So we need to refine this idea.

A better phrasing is that an offline state is “detached from its normal function” (Nichols et al., 1996, p.42).<sup>15</sup> But our question is what exactly this means: which aspects of a state's functional role is it detached from?

Consider mental imagery, which is sometimes characterised neurologically as “perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality.” (Nanay 2016a, p.66, cf. Kosslyn 1995, p.267). This is a difference on the input side, but an equally notable difference seems to be present on the output side: mental images don't lead to beliefs about our environment the way that perceptions do. Of course they can, with help from other states, contribute to more complex beliefs (e.g. visualising a scene from a different angle might lead to the belief about how it *would look*, given my beliefs about what's in the scene),

<sup>14</sup> There is some plausibility to the thought that in such cases the imagining can cause action only ‘indirectly’, or only in conjunction with other, non-imaginative, states (Funkhouser and Spaulding, 2009 defend this account of pretence) but then we need to spell out the relevant sort of indirectness.

<sup>15</sup> Notable terms in the same ballpark as ‘detaching’ include ‘quarantining’ (Goldman 2006a, 2006b), ‘backgrounding’ (Stokes, 2006), and ‘compartmentalising’ (Friend, 2020).

but they don't support beliefs in the direct way that perceptions do. However, mental images are not cut off from *i*-beliefs the way they are from beliefs: visualising something a certain way can be directly linked to a belief-like imagining that things are that way in a fictional world, a fantastic scenario, or some other imaginative project, so that imaginings within the same project interact the way that their online counterparts would. And this ability to interact with *i*-beliefs the way perceptions would interact with beliefs seems to be partly explained by mental images resembling perceptions in several ways that go beyond (but are plausibly explained by) a shared neural basis: their phenomenological quality,<sup>16</sup> the type of content they can represent, the way they are formatted and structured, and so on. Moreover, to the extent that mental images resemble perceptions in these ways, they can potentially generate most of the effects of perceptions outside their key evidential, belief-supporting, effect: if a perception of my cat could remind me that it's time to feed them, or evoke warm fuzzy associations, or let me estimate their tail-to-body ratio, then a matching mental image, if sufficiently vivid, can do all the same things.

We might try to sum this up by saying that mental imagery preserves most features of perception while removing (cancelling, suspending) its most tightly regulated functional properties, but retains those functional properties when interacting with other imaginings, as long as they are parts of the same imaginative project—the same fiction, the same thought experiment, the same simulated perspective, etc.

Consider next *i*-beliefs. They too can replicate many of the effects of belief, with the very notable exception that 1) they don't have to be formed in line with our evidence, and 2) they don't support believing their direct implications. Moreover, they can arise from and give support for other belief-like states just as a belief does, as long as those other states are themselves offline: that is, in fact, seemingly key both to the usefulness of instructive imagining, like counterfactual reasoning, and to the coherence of fictional worlds.<sup>17</sup>

It seems right to say roughly the same thing here as about mental imagery: *i*-belief preserves most features of belief while removing their mostly tightly regulated functional properties, but retains those functional properties when interacting with other imaginings.

<sup>16</sup> There is of course the common observation that sensory imaginings are typically 'less intense' or 'less vivid' in their phenomenology than their online counterparts. I think this is true as a tendency, just as it is also true that sensory imaginings are typically much less determinate in their content than their online counterparts. But it does not look like this is a matter of the two types of state varying within two markedly different ranges: rather, it is that perception is almost always close to the maximum of possible vividness and determinacy, while mental imagery varies widely, sometimes being incredibly faint and vague, sometimes virtually matching the vividness and determinacy of perception.

<sup>17</sup> People often distinguish, in reasoning about possibilities, between merely 'supposing' and actively 'imagining' (and/or 'conceiving'), where the former is a characteristically easier, 'thinner', activity than the latter. Explaining this distinction is a challenge for a unified recreativist account, since they have quite different properties despite both seeming to be in some sense belief-like (see, e.g., Kind, 2013, Balcerak Jackson, 2016, Arcangeli, 2018). My favoured view is that this distinction is not so much a matter of two distinct sorts of state, but of imaginative projects that activate multiple mental systems to enrich and unfold an imagined proposition, versus those that do not.

A final useful category to consider is motor imagery, the recreation of experiences of acting, which Nanay defines in a parallel way to mental imagery, as “processing in the motor cortex without bodily movement” (Nanay 2020, p. 394, cf. Jeannerod, 1994). Here again I think we can see the same pattern: motor imagery preserves most features of motor experience while removing the really key functional property, the direct connection to action.

This suggests a way to improve on the idea that offline states are ‘detached from their normal function’. It is not a state’s normal function *in general* that offlining detaches it from but, but certain specific components of it. This often leads to disconnection from external action, but that is neither necessary nor sufficient. So refining this common gloss, we might try to think of each type of offline state as being released from a kind of ‘core role’, a subset of its full causal profile, such that it only plays that core role in its interactions with other imaginings (and then only when the imaginings are bound together into a single imaginative project). But how, in general, do we tell what a state’s core role is?

What seems most notable to me about the functional properties that differentiate mental images, i-beliefs, and motor imagery differ from their online counterparts is that they are the sort of functions we habitually think of in terms of *normative constraints*. For a perception to not reflect environmental stimuli, for instance, is not just unusual but a sort of ‘malfunction’, as is a motor program which does not cause muscular movement; mental imagery and motor imagery differ precisely in that they can fail to have these connections, without that being in any way a defect. And for someone to not form beliefs in line with the evidence provided by their perceptions and other beliefs is not just an unusual variation in mental workings but a personal failure, an instance of irrationality.<sup>18</sup> Offline states are ‘liberated’ from evidential constraints, but conversely they also lose their power to directly affect the evidential constraints on other states. Other things that mental states may reliably do, like grabbing or directing attention, adjusting moods, or revealing geometric relations among shapes, are neither so central to their definition nor so norm-governed, and offline states can do those things just as much as online ones.

So I would suggest a sort of ‘normative constraint heuristic’. An offline version of a state differs from its online counterpart in being released from the latter’s ‘core role’, i.e. those functional respects that are subject to normative constraints. Note that this does not require core roles to be normative in any irreducible or mysterious sense. The sense in which perception ‘ought’ to be caused by incoming stimuli might simply be that perception being caused this way is a regularity in how our mind works that is so routine and so important that we cannot help seeing exceptions as defects. My use of normative language here is not intended to be committal

<sup>18</sup> Clearly there is a difference between the criticism of a subject as epistemically irrational and the criticism of a perception as malfunctioning. But what unites them is that they are inherent norms, norms rooted in the sort of state that is at issue—it is part of being a perception to be sensitive to stimuli, just as it is part of being a belief to be the sort of state that can warrant epistemic criticism. They contrast with external standards, like usefulness or pleasantness. An unwarranted belief might do very well by these standards: it might be exactly what we want to believe, might be useful, might make us feel good. None of this would erase its unwarrantedness.

about the ultimate validity of these norms, but to draw our attention to constraints on the workings of our mind which are especially deep-rooted and stable. It is a load-bearing feature of our mental architecture that perceptions are for the most part determined by incoming stimuli: we recognise this when we label deviations as malfunctions. My claim is that our ability to imagine comes from another feature of our mental architecture: that we can flexibly detach processes from these load-bearing constraints.

What this implies for *i*-desires is that whether a conative state is offline (an *i*-desire) or online (a desire), depends on whether it plays the core role of desire, i.e. the central and norm-governed parts of the ‘desire’ functional role.<sup>19</sup> So let us turn to the question of what that role is.

## 4 The Core Role of Desire

Let’s recap. I claim that it is both common and unproblematic for desire-like states to be indeterminately offline—to be halfway between desires and *i*-desires. This contrasts with the relationship between beliefs and *i*-beliefs (and mental/motor images and their online counterparts): they can usually be kept clearly apart, and it is an important part of responsible epistemic agency to make sure they are. The question is what explains this disanalogy. I’ve tried to motivate a certain way of thinking about offline states in general: that they are distinguished from their online counterparts primarily through the suspension of those parts of the latter’s functional role that are tightly regulated—specifically, the parts of that role that we naturally think about in terms of normative constraint. So we now have to ask what parts of the ‘desire’ role are thus regulated, and how they differ from the corresponding parts of the ‘belief’ role.

I think there are two key differences between desires and beliefs here, one about how they are formed and one about their effects. Both serve to bring the functional profile of an *i*-desire closer to that of a desire than an *i*-belief’s is to a belief’s.

### 4.1 The Causes of Desire

The first difference is that there is no required sort of cause for desires: although some desires are grounded in some other sort of state, some just appear, and there is generally nothing defective or problematic about them for it. This contrasts with beliefs: beliefs should be formed based on evidence of some kind, whether perceptual evidence, testimonial evidence, a compelling argument, or at minimum some sort of intuitive ‘seeming to be true’. If we learn that somebody believes something, it makes sense to ask them why—to expect something to ground and support that

<sup>19</sup> Note that by focusing on desire’s functional properties, and more specifically on its normatively-constrained functional properties, this approach thus evades the dilemma advanced by Tagliafico (2011), according to which a pretend desire must differ from a genuine desire either in its content or in the affect or emotion it involves.

belief. If nothing does, the baseless belief is shown to be problematic: the person should abandon it as far as they can. Not so with desires.

Of course it does sometimes make sense to interrogate a desire, to ask ‘why would you want *that*?’ But often this is out of curiosity: we assume that there probably *is* some further explanation, not that such an explanation is necessary for the desire’s legitimacy. An answer that says ‘no reason, that’s just what I want’ is perfectly respectable, even when it is surprising. There is nothing wrong with wanting something just because you want it; by contrast, believing something just because you believe it is a paradigmatic failure of epistemic rationality.<sup>20</sup>

Other times, we ask ‘why would you want that?’ because we see some obvious competing consideration, some reason not to desire that thing. Typically this has to be something stronger than just a reason why the actual fulfillment of the desire would be, in some respect, undesirable: after all, it is common and normal to have conflicting desires, and to therefore know that they cannot all be satisfied. As Kind puts it, “I might want to get some work done on Saturday and yet also want to spend the entire day with my children; I might want my elderly dog to live a long time and yet also want to put an end to his suffering” (Kind 2016-a, p.172). These pairs of desires cannot both be satisfied, but that does not make it irrational or otherwise problematic to feel them.

It might be claimed that just as beliefs constitutively aim at the true, desires constitutively aim at the good, so that a desire is defective unless its object is in some way good. A desire to consume something with no nutritional value and that gave no pleasure might then be thought a defective desire, analogous to a belief that had no supporting evidence. But we must be careful with the notion of ‘good’ here: on one reading, desires aim at the good simply in that desiring something constitutes ‘seeing it as good’, and that sort of norm is very different from the evidential norm on beliefs, since all our desires, as long as we have them, will present themselves as conforming to the norm. On the other hand, ‘good’ might mean something more substantive—e.g. an evolutionary function, or the promotion of pleasure, or some specific vision of human flourishing. Then a defective desire would be any desire for something to happen that would not, as far as the subject knows, contribute to the relevant substantive good (like the compulsion to turn on radios discussed by Quinn, 1993, pp.236–237, though see Smithies & Weiss, 2019 for an argument that this case is not really a desire, properly speaking). But even if some such substantive notion of ‘good’ was really a constitutive norm on desire, it arguably still does not play the same sort of role as the evidential norm of belief, for two reasons. First, there’s no consensus on what it is and how to apply it. Second, if the norm includes ‘bringing me pleasure’ as a way to be good (which seems the least controversial candidate good), then the fact that seeing any desire satisfied usually brings us a degree

<sup>20</sup> There is room for different theoretical explanations of this fact and its significance: that desires by themselves provide reasons, or that they create agent-relative reasons, or that they make it seem as if a reason exists, or that talk of practical reasons just is talk of desires (for discussion see, e.g., Darwall, 2001). All I need for my point here is the pretheoretical observation that ‘because I want to’ is often an adequate response to ‘why are you doing that?’ in a way that ‘because I think so’ is not an adequate response to ‘why do you think that?’.

of pleasure will mean that, just like on the thinner notion of ‘good’, any desire we have will conform to the norm simply in virtue of us having it.<sup>21</sup>

Sometimes it is held to be morally bad to desire evil things, such as the torture or murder of a rival, even if we recognise its evilness and so never act on those desires. I don’t think there is a decisive consensus on whether such desires are, in and of themselves, bad to have; for my own part I am inclined to think that moral evaluation should be confined to actions and beliefs. But either way, there remains a clear difference from beliefs: we need a positive reason to believe something, not just the absence of a reason not to believe it, but a desire is condemned only by some specific objection, not just by the absence of sufficient support.

This has important implications for i-desires. Part of what distinguishes i-beliefs from beliefs is that they do not require evidence, while beliefs do. We can imagine all sorts of things that go far beyond our evidence-based beliefs or even conflict with them, and it is crucial for this freedom that we are imagining, not believing. If they were beliefs, then we would need to evaluate their evidential basis, and reject any without sufficient support. But since they are mere i-beliefs, we need not. The result is that beliefs and i-beliefs behave very differently, and we are under rational pressure to keep the distinction clear.

But no parallel distinction holds between desires and i-desires. Neither requires a justification; while some desires might in fact follow from other mental states, like more general desires or moral principles, those that do not are not for that reason objectionable or irrational. So if we find, while watching a play or reading a novel, that a certain desire-like state arises in us with no obvious grounding in any pre-existing state, that fact is equally consistent with it being a desire or an i-desire. Just from looking at its origins, we have neither an obvious indication of its nature, nor a compelling reason to care.

That is not to say that there is no difference between the inputs to desires and i-desires. First, if a state clearly derives from a genuine desire (in conjunction with beliefs), that supports classifying it as itself a genuine desire, not an i-desire. Conversely, if a state clearly derives from an intention to simulate someone else’s perspective, that supports classifying it as an i-desire. The problem is that if a state just arises, as many conative states do, that does not allow us to classify it either way.

Second, if there are in fact moral or other substantive constraints on desires themselves, then on my account they would not apply to i-desires. And this seems to fit some of the most common cases where people seem to speak of i-desires: “I want X to happen, but of course I don’t *really* want it, that would be horrible...” The Tony Soprano case that Doggett and Egan discuss is a prime example: part of the motivation for insisting that we don’t *really* wish this violent man well is that such a desire might be morally inappropriate.<sup>22</sup> But of course, to the extent that it is open

<sup>21</sup> Katharina Anna Sodoma and an anonymous referee for *Erkenntnis* have my gratitude for pushing me on these points.

<sup>22</sup> This point is somewhat complicated by the fact that moral evaluability might be thought to attach, not to desire itself, but to the pleasure that results from it—standard examples in this debate tend to involve ‘taking pleasure in innocent suffering’. If, as I’ve suggested, pleasure can result from both desires and i-desires, any difference between the moral evaluability of the two would risk being ‘washed away’ by the shared moral evaluability of the pleasure they yield. (Again, for my own part I am sceptical that



to dispute whether desires are morally evaluable in themselves, any distinguishing of desires from i-desires on this basis will be likewise open to dispute.

## 4.2 The Effects of Desire

The second functional feature of desires that I think is important here is the holism of their effects. Desires cause things only in conjunction with other states, principally beliefs. Of course desires are not alone in this: many states, including beliefs themselves, depend on other states to determine their full effects. Indeed this is plausibly a general truth about the mind. But the effects of desires are, I argue, more thoroughly holistic than is true of other mental states. In particular, while beliefs and desires each depend on the other to have effects on behaviour, beliefs support other beliefs independently of desire, while desires yield other desires only in conjunction with beliefs.

For example, if I believed that on a distant island, scientists had successfully cloned dinosaurs, that belief would have a huge range of implications: about where the most biodiverse locations in the world are, about whether viable human cloning is possible, about how long DNA can be preserved in amber, and so on. These consequences flow out purely in the domain of belief; my desires have no role to play. But imagining this will not have these consequences—so we can distinguish a situation where I believe this from one where I merely imagine it by looking at whether there are impacts on, say, my beliefs about human cloning. Moreover, if I have a standing desire to speak the truth, other things being equal, then any belief of mine will tend to have some consequences for my actions, namely what I will say when asked about whether it is true.

By contrast, a desire-like state concerning something remote from our everyday lives may not have any clear or definite effects, and so may not give any good indication of its status as real or offline. For example, what consequences might a genuine desire for the raptors on that distant island to escape from a *T. rex* have? It won't affect my behaviour except in conjunction with some belief about how I might influence events on that island; it won't affect my emotions except in conjunction with some belief about whether it has been satisfied or not; it won't even give me derivative, instrumental desires without beliefs about what would contribute to its satisfaction. Of course in this case I don't have any of the relevant beliefs: I only have imaginings, i-beliefs. As a result, even from a real desire, I either won't get any of those consequences, or will get only versions that are themselves offline (instrumental i-desires), or characteristically muted and volatile forms of them (my affective reactions are typical of those I have to fictions, rather than to real life). That is, I get just the sort of consequences I would expect from an offline state—specifically, the same consequences I would expect from an i-desire with that content. Because the

---

Footnote 22 (continued)

either pleasure or desire in themselves warrant moral evaluation.) But this further blurring of the boundary between them is just more grist for my mill.

imagined scenario is remote from my possible actions, a real desire and an i-desire will have the same pattern of effects.

That is not to say there are no differences in output between desires and i-desires. When I have real beliefs about how my actions might affect the object of a desire, I have all the necessary ingredients for real motivations to action, ones that might even override competing motivations. If it were an i-desire instead, such motivations would not be forthcoming: I might act out my i-desires in a game of make-believe, but as soon as a genuine competing motivation intruded, the i-desire based motivation would (if I am being rational) yield to it. Suppose, for instance, someone wants to ‘get inside the head’ of a serial killer and predict their movements, and so imagines wanting to murder people. They can readily tell that they have a merely offline desire, because they have no tendency to act on that desire: it does nothing to motivate them outside of their imagining, and so we should say that they do not genuinely desire to murder people.

But desires directed specifically at fictional things (the main focus of our debate) do not let us differentiate them in this way. On topics where all of our beliefs are offline, i.e. mere belief-like imaginings, desires and desire-like imaginings will have the same pattern of effects. And since, as noted above, they can potentially have the same causes—they can both just arise without further justification—there may be no clear way to distinguish online from offline desires in such cases.

## 5 The Introspection Argument

How do the claims made so far address the Introspection Argument? I have argued that desires, unlike beliefs, are not required to be formed subject to certain constraints, and so i-desires cannot always be distinguished from desires by their origin. And desires, unlike beliefs, have no normatively constrained effects all by themselves, and so i-desires cannot always be distinguished from desires by their consequences. A spontaneously-felt desire regarding events remote from our real lives, and a spontaneously-felt i-desire about the same events, could thus be indistinguishable in their causes and effects. What should we say about conative states that occupy this confusing position? I think the answer will have to depend, in each case, on certain counterfactuals that may be hard to evaluate. In essence, we have to ask: if we *did* have beliefs about how to connect these conative states to our own actions, would we be motivated to act? If the fiction was real, and we had the opportunity to affect it, would we still feel that we wanted this? Or would our these seeming desires evaporate or transform upon collision with reality?

To use our running example, would a viewer of *The Sopranos*, who feels themselves wanting Tony Soprano to escape the police, still feel the same way if their imaginings about Soprano were replaced with beliefs? Perhaps we persuade them that *The Sopranos* is really a dramatised documentary about the Mafia in New Jersey, and that everyone Tony killed was a real person who really died. For many people, I think there would be a fairly clear answer: in this counterfactual situation, their desires would change. The neglect of morality they permitted themselves in fiction now seems unacceptable, their sympathy is withdrawn, and they sincerely hope the

bastard gets caught. The truth of counterfactuals like this, it seems to me, is a good basis for declaring these viewers to feel only an i-desire, a mere imagining, when they view the fiction. Conversely, some viewers might respond differently; perhaps for them the glamour of violent self-assertion is only intensified by reality, and they desire the mob boss's prosperity even more (there is no shortage of people willing to glamourise and enable Soprano's real-life counterparts, after all). Those viewers, it seems to me, genuinely desire that Tony Soprano do well: this is an online desire, despite being directed at a fictional object.

But there is a third class of viewers: those about whom it is simply indeterminate how they would feel if their imaginings became beliefs. To put it in Lewisian terms, possible worlds where they withdraw their sympathy from Tony, and those where they continue wishing him well, might be equally close to actuality. Perhaps it would depend on the context, or on how they were feeling at the time, or on details of exactly what content their beliefs had to be given to match their imaginings as closely as possible (e.g. is the guy on the screen an actor, or did they hire the real mob boss to act out his crimes?). Perhaps it would depend on how they saw others around them reacting. If there is no determinate answer to what effects their state would have under these counterfactual circumstances (perhaps because there is no determinate way to select the right counterfactual to consider), then I think we are forced to conclude that the state itself is neither determinately online nor determinately offline: it is neither a real desire nor a mere desire-like imagining, but something in between, a mental state in limbo.

With crime dramas the indeterminately offline desires might be a minority, but with more fantastical fictions they may be much more prevalent. In watching *The Expanse*, I find myself rooting for the inhabitants of the Asteroid Belt against the military of Mars; do I really want the Belters to win, or is this just an offline desire? Unless I can derive that desire from my genuine convictions and desires, we will have to look at its effects to decide, and it is likely to have no practical effects because I cannot do anything about this interplanetary war. We can try to ask about counterfactuals: *if* I believed it was real, what would I want then? But this counterfactual requires such a radical change to my beliefs that it seems impossible to evaluate. If I believed that Mars and the Belt were at war, what else would I believe? Which planet would I believe myself to live on? Who do my family support? Unless we have non-arbitrary ways to answer these questions, we will not be able to answer the original counterfactual question, and the desire will be indeterminately offline.

Another widespread source of indeterminacy will be conative states directed at particular fictional individuals. Suppose that in the course of reading about Bob and Jim, I come to sympathise more with Bob, and i-desire that he wins out in their rivalry. Would I desire this if I really believed in Bob and Jim? Well, it would depend who I felt more sympathy and rapport with, which would depend on how I got to know them—which would certainly be very different from the way one gets to know fictional characters.<sup>23</sup> But there is plausibly no determinate right way to fill

<sup>23</sup> Perhaps I might most closely approximate, with a real person, the way I know fictional characters, by believing that I have read a long factual narration of a real person's activities; but even this invites further questions, like whether I should reasonably trust that narration. After all, people are different when you meet than from how any story—whether written by them, or not—portrays them.

out the details of how I *would have to have* come to know them, *if* my i-beliefs about them were to be replaced with beliefs. And so if the online/offline status depends on how to fill that in, my state will not be determinately offline or online.

Note that this is not mere epistemic indeterminacy: it's not just that we don't know some fact. We might remain unable to answer these questions even if we, hypothetically, came to know every fact about my current brain processes and the brain processes I would undergo under various possible conditions. The point is that our concept of 'desire' is not an absolutely precise concept, and it does not by itself specify exactly which dispositions, relative to exactly which circumstances, are sufficient to make the difference between being a perhaps weak and inconstant, but genuine, desire, and not-quite-being a desire. Any sufficiently complex psychological kind will admit of borderline cases; it is just that with desires and i-desires, the borderline region is unusually large.

I think this provides an answer to the Introspection Argument. The reason we're not used to thinking in terms of desires and i-desires is that what we actually experience is a whole spectrum of online, offline, and indeterminately-offline conative states, whose different statuses can often only be determined by careful consideration of hypotheticals. Given this spectrum, a binary distinction between the categories of 'real desire' and 'mere imagining-desire' will be hard to apply, and yield relatively little insight. It is not really an improvement on vague language like 'well, I want it, but I don't *really* want it'. It is, I think, a very useful distinction for theorists, especially those concerned with how desire relates to a broadly recreativist picture of the imagination. But it is too blunt a tool to be of broader use.

## 6 The Accountability Argument

The above is my response to the Introspection Argument. What about the Accountability Argument? How much sense does it make to take offline desires, or even indeterminately offline desires, as revealing of our character?

Here, the causal origin of a conative state may be more important than its online/offline status. Some i-desires are produced by an intentional stipulation: I want to predict the serial killer's next move, so I take up their perspective, which involves i-desiring to murder people. This reveals little about me, since the source of the i-desire is so obviously artificial.<sup>24</sup> But most of the i-desires we experience in enjoying fiction don't arise like this: they arise spontaneously from the way we are 'struck' by the things we are imagining in other ways—the things we are visualising, i-believing, and so on. We picture the characters and follow along the events, and see how they move us, what they stir up in us, which way they incline us to hope things go. This is, significantly, the same basic process by which we form many

<sup>24</sup> The degree to which I struggle to 'put on' this i-desire, or find myself 'slipping naturally into it', might perhaps be revealing (cf. Bailey, 2021)—though what it reveals about my conative dispositions will be confounded with what it reveals about my facility at simulating other people's strange desires, or willingness to try. And it at least does not reveal that I secretly find something appealing about murder.

of our regular, online, desires: although we sometimes form them through practical reasoning, or through reflection on values and ideals, or through the simple impulse to remove displeasure or prolong pleasure, a good chunk of our desires just ‘bubble up’ as we move around the world and interact with various objects in it.

The natural thing for recreativists to say, it seems to me, is that the causal origin is the same in the real-world and fictional cases: both reflect the operation of what we might refer to as ‘our emotional systems’, but might just as informatively call our ‘heart’. These systems take in rich information about situations, provided either by beliefs, perceptions, and other online states, or by i-beliefs, mental images, and other offline states (or by a mixture of the two) and generates overall evaluations, and preferences for how the situations should change. These evaluations and preferences may be online, i.e. part of the mind’s ‘regular economy’, governed by evidential, and other, constraints, or they may be offline, i.e. part of an imaginative project that is partly or fully ‘detached’ from that economy. But because they are both generated by the same operations of the heart, they can both reveal the secrets of the heart. That’s why i-desires, despite being offline, can teach us important things about ourselves.

But we should be wary of over-stating this revealingness. Online and offline desires can spring from the same source, but they evolve under different pressures. Most of us, if we had a friend who had killed several people, would experience a lot of conflicting and disorienting feelings and desires, which would prompt a lot of what we might call ‘soul-searching’. As a result of that reflection, we would probably try to exercise some degree of top-down control, regulating some desires in light of others. For instance, we might decide to actively try to limit our sympathy with this friend, and the concomitant desire for their well-being, and/or to actively to increase our sympathy for their victims, their potential future victims, and their victim’s loved ones. The desires we ended up having would then be doubly revealing: they show not just what sorts of desires spring up spontaneously from our heart, but also whether, how, and how effectively, we regulate them.

When engaging with fiction, this sort of soul-searching and self-regulation are less common. They are not always completely absent: part of the value of *The Sopranos* is precisely that it encourages this sort of reflection, that it simultaneously works to evoke sympathy with Tony Soprano and then to confront us with the moral reality of his actions. To the extent that we engage reflectively we are more likely to have either definitely online or definitely offline desires: we are prompted to confront, even if implicitly, the sort of counterfactual questions that matter for that distinction, about how we would feel if this were actually happening, if we were really involved, if we knew one or more the people affected, and so on. By considering the questions, we can solidify our dispositions, thereby making it more firmly the case that we either would or would not feel differently if the fiction were reality, and thereby making the ‘desires’ we feel more firmly either really online or merely offline.

But lots of TV shows go the opposite way: they actively try to relieve us of any pressure to reflect on these questions. For instance, arguably the majority of protagonists in certain genres (action, fantasy, scifi, thrillers) are responsible for multiple deaths, but the show goes out of its way to minimise their significance—the bad

guys wear face-concealing masks, or we don't get confirmation that they're dead but only that they've been blasted offscreen, or we see massive explosions and collapsing buildings without ever actually see any bodies. The grieving friends and family of henchman #7 never appear, and the police and prison guards to whom Spiderman hands over 'criminals' are never abusive or corrupt. While some shows make a point of presenting characters with difficult moral dilemmas, these ones make a point of keeping all their moral decisions simple and straightforward. The effect of these contrivances is to remove any pressure to render conative states definitely online or definitely offline. Do we really want the hero to succeed at blowing up the villain's compound? If we grappled with the fact that it's full of support staff who might be complicit but don't deserve death, we might experience a lot of moral anxiety and cognitive strain: but we're not watching this show in order to experience those things, so we don't grapple with them. The writers want us to keep watching, so they make that non-grappling easy.

In effect, I am suggesting that the limbo zone between desire and i-desire is deliberately cultivated as part of many aesthetic endeavours: many fictions aim specifically to arouse and gratify conative states, while avoiding the sort of engagement that would force us to decide what we really want, i.e. force them into being either determinately online or determinately offline. Sometimes we specifically want to not have to ask certain questions. And, to be clear, I think that is perfectly legitimate: there's no reason why we should be grappling with complexity all the time. Sometimes fiction can just be fun. And I think something stronger can be claimed: enjoying fiction that's just fun may be the best way to enable productive reflection. Part of the power of *The Sopranos* is that we can often recognise, in our disposition to sympathise with a charismatic murderer, some of the same dispositions that are operative in letting us enjoy the exploits of innumerable other protagonists, whose often larger body counts are more casually swept aside. We have more data to work with, in understanding our own hearts, if we sometimes let our hearts indulge themselves freely without constant scrutiny, rather than enforcing standards of realism, morality, and so on, on their every move. I think recognising not only desires and i-desires, but the broad spectrum of states in limbo between the two, helps to capture this complex sort of revealingness.

## 7 Conclusions

My defence of i-desires has been limited in a number of ways. I have not engaged with debates about their content, their powers to justify or motivate in simulation or pretence, or their relation to imaginative resistance or the supposing/conceiving distinction. And my defence is conditional on the viability of a broader recreativism about the imagination generally. What I have tried to show is that recreativism is well-placed to answer the Introspection Argument and the Accountability Argument, and that its answers are not ad hoc but reflect something significant about the nature of desire itself. Both of these arguments focus on ways that putative i-desires seem to differ importantly from other imaginative states: they are problematic for recreativism because it aims to offer an account that unifies them all. I have

argued that what is common to all imaginative states—their disengagement from a core role—will apply differently to states with different functional roles, in a way that yields precisely the pattern of differences that these anti-recreativist arguments appeal to. If I am right, then recreativism’s unifying ambitions are correspondingly strengthened.

## References

- Arcangeli, M. (2018). *Supposition and the imaginative realm: A philosophical inquiry*. Routledge.
- Bailey, O. (2021). *Empathy with vicious perspectives? A puzzle about the moral limits of empathetic imagination*. Synthese.
- Carroll, N. (1990). *The philosophy of horror, or, paradoxes of the heart*. Routledge.
- Chalmers, D. (2002). Does conceivability entail possibility? In T. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 145–200). Oxford University Press.
- Currie, G. (1995). Imagination and simulation: Aesthetics meets cognitive science. In M. Davies & T. Stone (Eds.), *Mental simulation* (pp. 151–169). Blackwell.
- Currie, G. (1997). The paradox of caring: Fiction and the philosophy of mind. In M. Hjort & S. Laverac (Eds.), *Emotion and the arts* (pp. 63–77). Oxford University Press.
- Currie, G. (2002). Desire in imagination. In T. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 201–221). Oxford University Press.
- Currie, G. (2010). Tragedy. *Analysis*, 70(4), 632–638.
- Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. Clarendon Press.
- Darwall, S. (2001). Because i want it. *Journal of Social Philosophy and Policy*, 18, 129–153.
- Doggett, T., & Egan, A. (2007). Wanting things you don’t want: The case for an imaginative analogue of desire. *Philosophers’ Imprint*, 7(9), 1–17.
- Doggett, T., & Egan, A. (2012). How we feel about terrible, non-existent Mafiosi. *Philosophy and Phenomenological Research*, 84(2), 277–306.
- Dorsch, F. (2011). Emotional Imagining and Our Responses to Fiction. *Enrahonar: Quaderns de Filosofia*, 46, 153–176.
- Dorsch, F. (2012). *The unity of imagining*. De Gruyter.
- Dorsch, F. (2016). Knowledge by imagination: How imaginative experience can ground knowledge. *Teorema: International Journal of Philosophy*, 35(3), 87–116.
- Egan, A. (2009). Imagination, delusion, and self-deception. In T. Bayne & J. Fernández (Eds.), *Delusions and self-deception: Motivational and affective influences on belief formation* (pp. 263–280). Psychology Press.
- Friend, S. (2003). How i really feel about JFK. In M. Kieran & D. Lopes (Eds.), *Imagination, philosophy, and the arts* (pp. 35–54). Routledge.
- Friend, S. (2020). Fiction and emotion: The puzzle of divergent norms. *The British Journal of Aesthetics*, 60(4), 403–418.
- Funkhouser, E., & Spaulding, S. (2009). Imagination and other scripts. *Philosophical Studies*, 143(3), 291–314.
- Gaut, B. (2003). Reasons, emotions, and fictions. In M. Kieran & D. Lopes (Eds.), *Imagination, philosophy, and the arts* (pp. 14–34). Routledge.
- Gendler, T. S. (2000). The puzzle of imaginative resistance. *Journal of Philosophy*, 97(2), 55–81.
- Goldman, A. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press.
- Goldman, A. (2006). Imagination and simulation in audience responses to fiction. In S. Nichols (Ed.), *The architecture of the imagination: New essays on pretence, possibility, and fiction* (pp. 41–56). Clarendon Press.
- Gordon, R. (1995). Simulation without introspection or inference from me to you. In M. Davies & T. Stone (Eds.), *Mental simulation* (pp. 352–366). Blackwell.
- Hazlett, A. (2009). How to defend response moralism. *British Journal of Aesthetics*, 49(3), 241–255.
- Heal, J. (1998). Co-cognition and off-line simulation. *Mind and Language*, 13, 477–498.

- Jackson, M. (2016). Imagining, supposing and conceiving. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 41–60). Oxford University Press.
- Jackson, M. (2018). Justification by imagination. In F. Dorsch & F. Macpherson (Eds.), *Perceptual memory and perceptual imagination* (pp. 209–226). Oxford University Press.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–245.
- Kind, A. (2011). The puzzle of imaginative desire. *Australasian Journal of Philosophy*, 89, 421–439.
- Kind, A. (2013). The heterogeneity of the imagination. *Erkenntnis*, 78(1), 141–159.
- Kind, A. (2016). Desire-like imagination. In A. Kind (Ed.), *Routledge handbook of philosophy of imagination* (pp. 163–176). Routledge.
- Kind, A. (2016). Imagining under constraints. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 145–159). Oxford University Press.
- Kind, A. (2018). How imagination gives rise to knowledge. In F. Dorsch & F. Macpherson (Eds.), *Perceptual memory and perceptual imagination* (pp. 227–246). Oxford University Press.
- Kind, A., & Kung, P. (2016). Introduction: The puzzle of imaginative use. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 1–37). Oxford University Press.
- Kung, P. (2010). Imagining as a guide to possibility. *Philosophy and Phenomenological Research*, 81(3), 620–633.
- Langland-Hassan, P. (2020). *Explaining imagination*. Oxford University Press.
- Liao, S.-y, Strohminger, N., & Sripada, C. (2014). Empirically investigating imaginative resistance. *British Journal of Aesthetics*, 54(3), 339–355.
- Myers, Joshua. (2021a). Reasoning with imagination. In A. Kind & C. Badura (Eds.), *Epistemic uses of the imagination*. Routledge.
- Myers, Joshua. (2021b). The epistemic status of the imagination. *Philosophical Studies*, 178(10), 3251–3270.
- Nanay, B. (2016). Hallucination as mental imagery. *Journal of Consciousness Studies*, 23(7–8), 65–81.
- Nanay, B. (2016). The role of imagination in decision-making. *Mind and Language*, 31(1), 126–142.
- Nanay, B. (2020). Motor imagery and action execution. *Ergo: An Open Access Journal of Philosophy*, 7(13), 391–408.
- Nichols, S. (2004). Review of currie and ravenscroft, recreative minds. *Mind*, 113, 329–334.
- Nichols, S., & Stich, S. (2003). *Mindreading*. Oxford University Press.
- Nichols, S., Stich, S., Leslie, A., & Klein, D. (1996). Varieties of off-line simulation. In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind* (pp. 39–74). Cambridge University Press.
- Peterson, E. (2021). Attention to details: Imagination, attention, and epistemic significance. In A. Kind & C. Badura (Eds.), *Epistemic uses of the imagination*. Routledge.
- Quinn, W. (1993). Putting rationality in its place. In P. Foot (Ed.), *Morality and Action* (pp. 228–255). Cambridge, UK: Cambridge University Press.
- Radford, C. (1975). How can we be moved by the fate of Anna Karenina? *Proceedings of the Aristotelian Society, Supplemental*, 49, 67–80.
- Schellenberg, S. (2013). Belief and desire in imagination and immersion. *The Journal of Philosophy*, 110(9), 497–517.
- Schwitzgebel, E. (2001). In-between believing. *Philosophical Quarterly*, 51, 76–82.
- Smithies, D., & Weiss, J. (2019). Affective experience, desire, and reasons for action. *Analytic Philosophy*, 60(1), 27–54.
- Smuts, A. (2015). How not to defend response moralism. *Journal of Aesthetic Education*, 49(4), 19–38.
- Smuts, A. (2016). The ethics of imagination and fantasy. In A. Kind (Ed.), *The Routledge handbook of philosophy of imagination* (pp. 380–391). Routledge.
- Spaulding, S. (2015). Imagination, desire, and rationality. *Journal of Philosophy*, 112(9), 457–476.
- Stear, N.-H. (2009). Sadomasochism as make-believe. *Hypatia*, 24(2), 21–38.
- Steuber, K. (2016). Empathy and the imagination. In A. Kind (Ed.), *The Routledge handbook of philosophy of imagination* (pp. 368–379). UK: Routledge.
- Stokes, D. (2006). On the evaluative character of imaginative resistance. *British Journal of Aesthetics*, 46, 207–245.
- Tagliafico, D. (2011). Can we really speak of ‘pretend desire’? *Logic and Philosophy of Science*, 9(1), 461–467.
- van Leeuwen, N. (2016). The imaginative agent. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 85–109). Oxford University Press.



- van Leeuwen, P. (2016). Imagination and action. In A. Kind (Ed.), *Routledge handbook of philosophy of imagination* (pp. 286–299). Routledge.
- Velleman, David. (2000). On the Aim of Belief. In D. Velleman (Ed.), *The Possibility of Practical Reason* (pp. 244–281). New York: Oxford University Press.
- Walton, K. (1978). Fearing fictions. *Journal of Philosophy*, 75(1), 5–27.
- Walton, K. (1990). *Mimesis as make-believe*. Harvard University Press.
- Walton, K. (1997). Spelunking, simulation, and slime: On being moved by fiction. In M. Hjort & S. Laverac (Eds.), *Emotion and the arts* (pp. 37–49). Oxford University Press.
- Weinberg, J., & Meskin, A. (2005). Imagine that. In M. Kieran (Ed.), *Contemporary debates in aesthetics and the philosophy of art*. Blackwell.
- Williamson, T. (2016). Knowing by imagining. In A. Kind & P. Kung (Eds.), *Knowledge through imagination* (pp. 113–123). Oxford University Press.
- Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, 53(1), 1–42.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.