

This article appeared in *The Journal of Aesthetics and Art Criticism*, Special issue on Art, Mind, and Cognitive Science, 2004.

SHAUN NICHOLS

Imagining and Believing: The Promise of a Single Code

Abstract:

Recent cognitive accounts of the imagination propose that imagining and believing are in the same “code”. According to the single code hypothesis, cognitive mechanisms that can take input from both imagining and from believing will process imagination-based inputs (“pretense representations”) and isomorphic beliefs in much the same way. In this paper, I argue that the single code hypothesis provides a unified and independently motivated explanation for a wide range of puzzles surrounding fiction.

The imagination has always been one of the darker faculties of the human mind. Recent work on the imagination in cognitive science scarcely dispels all the mysteries of the imagination. But cognitive approaches might illuminate one central feature of imagination – the striking similarities between imagining and believing. In this paper I will explore the idea that the similarities between imagination and belief are partly explained by the hypothesis that imagination and belief are in a “single code”. A good deal of scaffolding is necessary to explain this idea, but I will argue that the single code hypothesis provides a surprisingly powerful account of one aspect of the imagination.

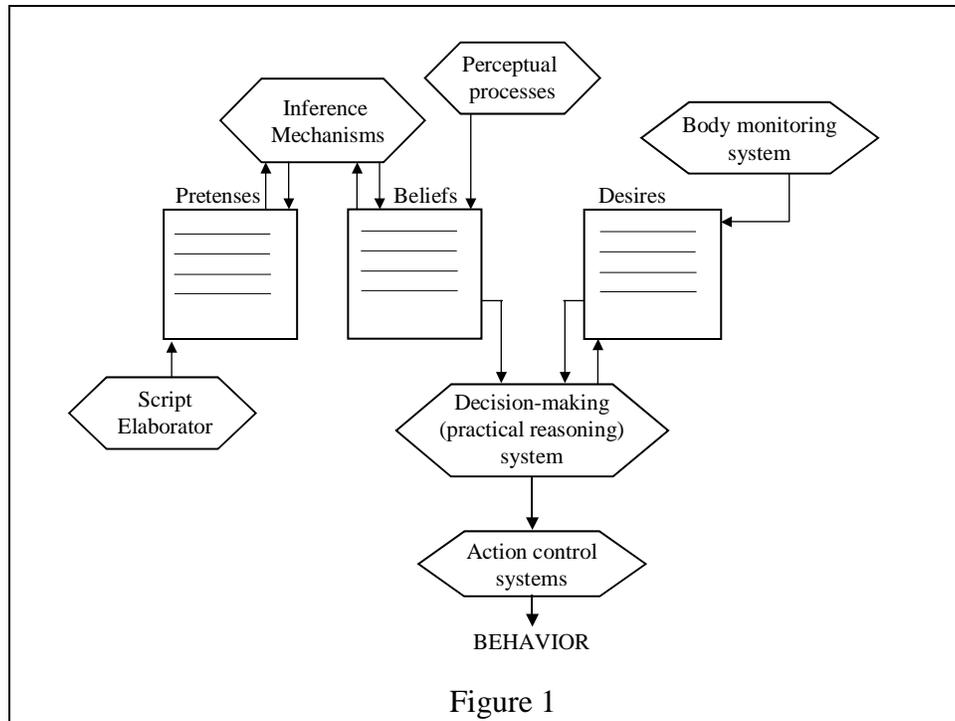
Unraveling the details of the relation between imagining and believing has ascended to a prominent place on the agenda of researchers working on theories of imagination, fiction, and belief. In this paper, I will not be defending an underdog theory. Far from it – the single code theory, in one form or another, is embraced by most contemporary cognitivist accounts of pretense and imagination.¹ However, the explanatory power of the single code hypothesis has not, I’ll maintain, been adequately recognized. The single code hypothesis provides a unified explanation for a wide range of similarities between imagination and belief. In the first section of this paper, I will describe and elaborate the version of the single code theory that Stephen Stich and I have developed. Then I will argue that this theory offers simple solutions to a diverse set of philosophical puzzles about fiction.²

I. FICTION AND IMAGINATION

In earlier work, Stephen Stich and I developed a cognitive theory of the imagination, based largely on empirical findings and empirical considerations about pretend play.³ A brief review of a portion of our theory of pretend play will supply backdrop. Several features of the theory can be illustrated with a lovely experiment on children by Alan Leslie. In Leslie’s experiment, the child pretends to fill two cups with tea. The experimenter picks up one of the cups, upends it, and places it next to the other cup. Then the experimenter asks the child to point at the “full cup” and at the “empty cup”. Both cups are really empty throughout the entire procedure, but two-year-olds reliably indicate that the “empty cup” is the one that had been turned upside down and the “full cup” is the other one.⁴

One significant feature of episodes like Leslie's tea party is that children distinguish what is pretend from what is real. That is, at no point in this experiment do children believe that either of the cups is full. Their pretense that the cup is full is "quarantined" from their belief that the cup is not full. Another, more interesting, feature of the experiment is that it indicates that a belief and a pretense can have exactly the same *content*. When the children are asked to point to the "empty cup" and the "full cup", they maintain that the previously overturned cup is the empty one. On the most natural interpretation of this, the child is *pretending that the cup is empty*. We adopt the representationalist approach that is common in this area and say that such pretending involves a "pretense representation" with the content *the cup is empty*. Although the child is pretending that the cup is empty, she is not blind to the fact that the cup is really empty throughout; rather, the child also *believes that the cup is empty*. This suggests that the crucial difference between "pretense representations" and beliefs is not given by the *content* of the representation. For a pretense representation and a belief can have exactly the same content. So, pretense representations are quarantined from beliefs, and yet the distinction is not driven by differences in content. The natural cognitivist proposal, then, is that pretense representations differ from belief representations by their *function*. Just as desires are distinguished from beliefs by their characteristic functional roles, so too pretenses are distinguished from beliefs.⁵ Stich and I exploit the familiar illustrative device of using boxes to represent functional groupings, and we propose that besides a belief box and a desire box, there is a "pretense box".⁶ This pretense box, we suggest, is part of the basic architecture of the human mind.

In addition to a pretense box, Stich and I propose a mechanism that supplies the pretense box with representations that initiate or embellish an episode of pretense, the "Script Elaborator". This is required to explain the bizarre and creative elements that are evident in much pretend play. However, there are also much more staid and predictable elaborations in pretend play. This too is well illustrated by Leslie's experiment. Virtually all of the children in his experiment responded the same way when asked to point to the "empty cup". How are these orderly patterns to be explained? In everyday life when we acquire new beliefs, we routinely draw inferences and update our beliefs. No one knows how this process works, but no one disputes that it does work. There must be some set of mechanisms subserving inference and updating, and we can simply use another functional grouping to collect these mechanisms under the heading "Inference Mechanisms". Now, to explain the orderly responses of the children in Leslie's experiment, we propose that the representations in the pretense box are processed by the same inference mechanisms that operate over real beliefs. Of course, to draw these inferences the child must be able to use real world knowledge about the effects of gravity and so forth, and so Stich and I also suppose that the inferences the child makes during pretense can somehow draw on the child's beliefs.⁷ Putting this all together generates the functional architecture depicted in Figure 1.



With the functional architecture in place, we can finally articulate the core idea that I want to explore in this paper. Stich and I suggest that the representations in the pretense box and the representations in the belief box are in the “same code”, to adopt a computational locution. Of course, it’s far from clear what the code is for representations in the belief box, so it’s not possible to be specific about the details or the nature of the putatively shared code.⁸ But the important point for present purposes can be made without giving further detail about what the code is. If pretense representations and beliefs are in the same code, then mechanisms that take input from the pretense box and from the belief box will treat parallel representations much the same way.⁹ For instance, if a mechanism takes pretense representations as input, the single code hypothesis maintains that if that mechanism is *activated* by the occurrent belief that *p*, it will also be activated by the occurrent pretense representation that *p*. More generally, for any mechanism that takes input from both the pretense box and the belief box, the pretense representation *p* will be processed much the same way as the belief representation *p*. I will count any theory that makes this claim as a “single code” theory.

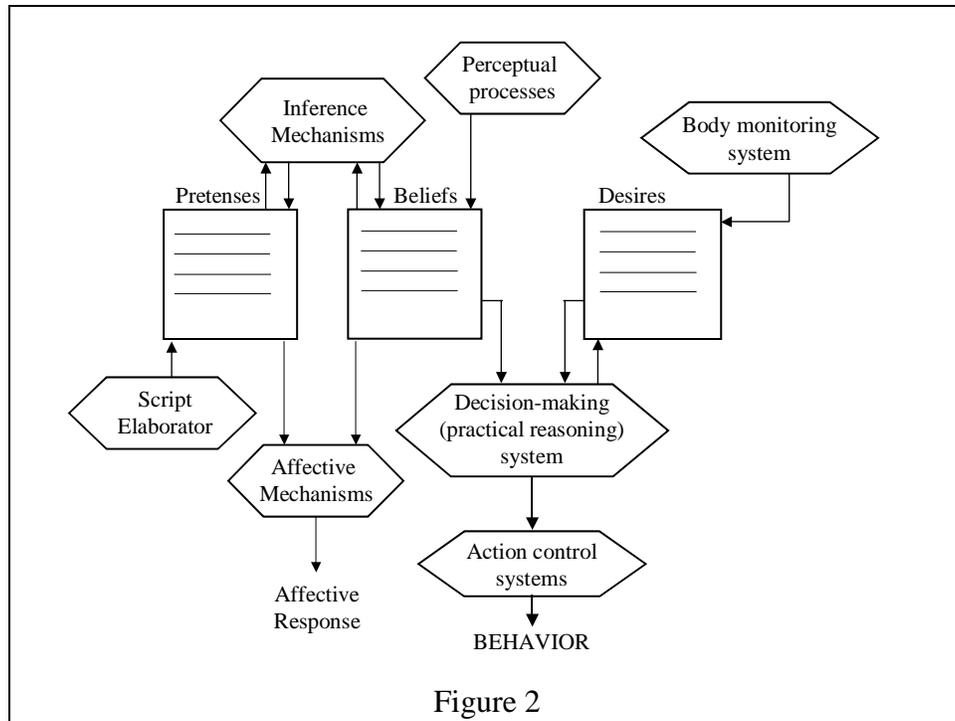
Taking a cognitive approach to imagination, it is worth emphasizing, provides the crucial resources for articulating the single code proposal. For the single code hypothesis is a proposal about the relation between certain kinds of representations and certain kinds of cognitive mechanisms. As a result, the hypothesis can only be framed against a background of cognitive architecture. Once one has posited a background of cognitive components, one can then explain some processing differences between imagination and belief by noting that some of these cognitive components take input from beliefs but not from imagination. The single code hypothesis maintains that one can also explain many of the similarities between imagination and belief with the proposal that some of the cognitive mechanisms do process input from both

imagination and belief, and further, that pretense representations are in the same code as belief representations.

The core claim of the single code hypothesis is that a mechanism that takes pretense representations as input will process those representations much as it would process isomorphic belief representations. As a result, off-line simulation theorists are plausibly viewed as single code theorists. For they maintain that several mental mechanisms process ‘pretend beliefs’ just like real beliefs.¹⁰ Off-line simulation theorists often have additional commitments of course. For instance, many prominent versions of off-line simulation theory explicitly invoke pretend desires in addition to pretend beliefs, and also maintain that the practical reasoning system takes as input pretend beliefs and pretend desires. Those additional stipulations are consistent with, but not required by, the single code hypothesis.¹¹ Although off-line simulation theory counts as a single code theory, the development of the idea in this paper will lean on the account of pretense that Stich and I developed (the “pretense box” account) because it will be much easier to investigate these issues with a detailed account in hand.

The pretense box account was developed as an account of pretend play, but we also maintained that the account provided a theory of the imagination.¹² If this account of the imagination is right, it should also accommodate imaginative engagement with fiction. The extension to fiction is obvious enough. In consuming fiction, one (or one’s cognitive mechanisms) constructs a set of representations in one’s pretense box. This construction is guided by a prompt of some sort; the prompt might be a movie, a novel, or a raconteur. The representations being put into the pretense box are distinguished from beliefs by their functional roles, but they are not systematically distinguished from beliefs in terms of content. As with imagination generally, the pretense representations generated by fiction are in the same mental code as beliefs. In consuming fiction, then, the inferential systems process the pretense representations much as the inferential systems would process isomorphic beliefs.¹³

On the single code hypothesis, if a mechanism takes pretense representations as input, that mechanism will process the pretense representation much the same way it would process an isomorphic belief. This should apply to emotional systems as well as inferential systems. According to the single code hypothesis, then, the emotional systems will respond to pretense representations much as they do to parallel beliefs.¹⁴ That is, if the pretense representation that *p* gets processed by an affective mechanism, the affective outputs should parallel those of the belief that *p*. This claim gains support from a wide range of empirical findings. Imagining that something dangerous is happening can provoke affective response that at least closely resembles real fear. In one representative experiment, subjects were presented with an imaginary scenario involving a snake encounter, and virtually all of the subjects showed physiological signs associated with fear. More interestingly, the responses were stronger in people who were snake phobic than in people who were phobic about public speaking.¹⁵ So the particular character of the subject’s affective response system seems to be reflected in their affective responses to imagination. All of this suggests that we need to elaborate the functional architecture to capture the fact that affective response systems can take input from both beliefs and pretense representations, hence Figure 2.¹⁶



It remains an open question on the single code hypothesis whether all affective systems take pretense representations as input. But we might as well begin with the bold supposition that all affective and emotional systems that take beliefs as input also take pretense representations as input. It might turn out, however, that certain affective or emotional systems will take beliefs as inputs but not pretense representations. As philosophers like to say, that's an empirical matter.¹⁷

II. PUZZLES IN FICTION

The pretense box account is driven by broadly empirical concerns, rather than by any considerations about fiction. But the account, if right, should accommodate imaginative engagement with fiction. So, how does the account fare with fiction? Much work in philosophy is driven by apparent puzzles, and this approach has been especially central in the philosophy of fiction. Indeed, Kendall Walton's important account of how we appreciate representational art is supposed to be measured largely by its success at solving puzzles surrounding fiction.¹⁸ In this section, I'll argue that the cognitive account of the imagination sketched above, and in particular, the single code hypothesis, provides solutions to a wide range of the puzzles surrounding fiction.

Before we start, however, a caveat is in order. Each of the puzzles that I describe has attracted significant theoretical attention, and for most of the puzzles there are a number of competing theories about how to solve the puzzle.¹⁹ For the most part, however, I won't engage alternative explanations of the puzzles, which will make the discussion rather breezy. This is because I'm not primarily trying to score points against other treatments of the puzzles; rather, the goal is to indicate the explanatory sweep of the single code hypothesis. The single code hypothesis is independently motivated by broadly empirical considerations, and I want to show that, across a wide range of puzzles, the single code hypothesis provides natural solutions.

i. Emotions and fiction

I begin with perhaps the most familiar puzzle, our apparent emotional responses to fiction. One way to put the puzzle is as an inconsistent triad:

I pity Lear.

To pity something, you have to believe it exists.

I don't believe Lear exists.²⁰

An important part of the puzzle here is that we know that Lear doesn't exist and yet we have an emotional response that seems to fit the emotional response that we would have if we did believe that Lear existed. The puzzling phenomenon here holds across a wide range of emotions. We respond to fiction with affect that closely parallels our affective responses to the real world. Novels and movies make us feel sad, happy, triumphant, indignant, embarrassed, afraid.

The single code hypothesis applies itself in an obvious way to the puzzle outlined, and other philosophers have ably promoted the virtues of single code accounts here, so I will be brief.²¹ Since the pretense representations generated by engaging with the fiction are in the same code as beliefs, these pretense representations will be processed by affective mechanisms much the way isomorphic beliefs would be processed. Single code theories thus promise a solution to at least the psychological puzzle of how fiction can cause affective response. Affective systems can receive input from the imagination, and affective systems process input from the imagination as they would process isomorphic beliefs. So on the single code theory, an affective system can accept as input the pretense representation that *Lear is watching Cordelia die*, and the affective system will generate the same affective consequence from this input it would if the input were instead a belief that *Lear is watching Cordelia die*. Since the 'pity'-response would be activated by the belief, it is activated by the isomorphic pretense representation. The single code hypothesis also explains why fiction can produce such a vast range of emotional responses. By hypothesis, affective systems that are triggered by beliefs can also be triggered by the imagination. Hence, our emotional reactions to imagining should be as diverse as our emotional reactions to believing.

ii. Emotion and iteration

An additional puzzle about emotion arises in the context of iterated pretense. Walton describes the puzzle as follows: "Consider a story that ends with the hero waking from a bad dream. The reader, on realizing that it is fictional (in the story) only that it is *fictional* (in a dream) that monsters were chasing him, not that they really were, heaves a sigh of (fictional) relief."²² One might quibble with this description of the puzzle, but the phenomenon is clear and interesting. When we read a novel in which something is merely pretended, the fact that it's pretend in the novel mitigates our emotional reaction. We are relieved when it is only pretend that the hero is harmed – as when we learn that the events were merely fantasies of the villain. But this seems odd – the *whole thing* is fiction.

The single code hypothesis promises to explain this puzzle as well. When we come to believe that some danger to another person was merely imaginary, we "heave a sigh of relief." Our affective response to the danger cues is mitigated. The single code hypothesis suggests that something parallel happens when we imagine that the danger to the fictional hero is only imaginary.²³ The parallel can be depicted more concretely as follows. The belief that Joe is dying creates an emotional response that is then mitigated by the subsequent belief that it is only fictional that Joe is dying; similarly, the pretense representation that Joe is dying creates an

emotional response that is mitigated by the subsequent pretense representation that it is only fictional that Joe is dying.²⁴

iii. Fictional names

The next puzzle that I want to consider has been set most fully by Gregory Currie. Most people do not believe that fictional characters really exist. Yet our intuitions about names of fictional characters parallel in a striking way our intuitions about names of real individuals. This is particularly evident in our reference intuitions for fictional names. We know that “Macbeth” does not refer to a concrete individual, and the play provides the only description of the character. As a result, perhaps the most obvious story about the semantics of fictional names is a straightforward description theory according to which the meaning of a fictional name is given by its associated description. However, a description theory of fictional names turns out to be quite counterintuitive. Indeed, there are closely parallel antidescriptivist intuitions regarding fictional proper names and actual proper names. These puzzling parallels demand some explanation, and the single code hypothesis provides one. To begin, I want to catalogue explicitly some of the parallel anti-descriptivist intuitions that are generated for fictional names and proper names.

Uniqueness

We have the intuition that fictional names refer to unique individuals, even when the description associated with the fictional name clearly does not pick out a unique individual. Currie concocts the following (very brief) story:

“Jack got up in the morning and ate breakfast.”

As Currie notes, we can’t simply gloss this as saying, “There is someone called ‘Jack’ who got up in the morning and ate breakfast,” since “the story purports to describe the activities of a particular Jack.”²⁵ Currie’s example here bears a close analogy to Kripke’s “Feynman” example against the description theory of proper names.²⁶ Someone might believe that Feynman was a physicist and not have any other distinguishing beliefs about Feynman. Of course, this description, *he was a physicist*, hardly picks out anyone uniquely. Yet we’re still inclined to say that the person is referring to a particular individual. Currie’s story of Jack similarly seems to pick out a single “fictional character,” despite the impoverished description.

Meeting the description isn’t sufficient

Even if there is a unique individual who actually meets the description associated with a fictional name, we have the intuition that the fictional name does not refer to that individual. Kripke calls attention to this fact when he claims that even if there is an actual person who happens to meet the description associated with Sherlock Holmes, the stories are not about this man.²⁷ Although it goes unremarked in *Naming and Necessity*, the Holmes example parallels Kripke’s Gödel/Schmidt case.²⁸ Kripke notes that the one thing that is commonly associated with Gödel is that he discovered the incompleteness of arithmetic. However, Kripke maintains, if it turned out that a man named “Schmidt” actually discovered the incompleteness of arithmetic, we would not assume that “Gödel” really referred to Schmidt. We have the intuition then, that meeting the description associated with “Gödel” does not suffice to be the referent of “Gödel”. Again, this parallels the intuition that meeting the description associated with “Sherlock Holmes” does not suffice to be the referent of “Sherlock Holmes”.

Meeting the description isn't necessary

When reading the Holmes stories, it seems that we want to say that Holmes might not have done some of the things that he did do – we don't regard the description associated with "Holmes" as necessary. Currie writes, "It is true in the story that if Holmes had not done the things described in the story he would still have been Holmes."²⁹ When reading the stories, it seems possible that lots of the descriptions that are attached to Holmes are accidental. For instance, it seems possible that in the story, Watson would discover that Moriarty had disguised himself as Holmes in a number of the adventures, while Holmes was tied up in Moriarty's den. Kripke makes a similar point about proper names: "Most of the things commonly attributed to Aristotle are things that Aristotle might not have done at all. In a situation in which he didn't do them, we would describe that as a situation in which *Aristotle* didn't do them."³⁰ So, just as we have the intuition that the bearer of a name needn't have the properties associated with that name, we also have the intuition when engaged with fiction that the bearer of the fictional name needn't have the properties associated with that (fictional) name.

The puzzle, then, is that we know that the fictional names don't refer to anything, yet our intuitions run against a descriptivist treatment of fictional names. Rather, the anti-descriptivist intuitions about proper names find close parallels in fictional names. Kripkean thought experiments seem to work equally well whether one uses a fictional name or a proper name to drive the intuition. The single code hypothesis can provide an explanation for the parallel intuitions.

Our goal here is not, I should stress, to determine the right semantics for fictional names. Rather, the goal is only to explain why we have these parallel intuitions about fictional names and proper names. To appreciate how the single code account explains the parallel, it's helpful to consider a related proposal. One idea for explaining our intuitions about fictional names is to maintain that these intuitions derive from metalinguistic pretenses; we pretend, for instance, that "*Sherlock Holmes*" refers or that "*Sherlock Holmes*" is a proper name.³¹ On this kind of proposal, the parallel intuitions follow from the fact that we pretend that the fictional names have the properties of actual proper names. On the single code hypothesis, by contrast, we can explain the intuitions without appealing to any explicit pretense that the fictional names refer. Since the representational code for pretense representations is the same as that for belief representations, the code for fictional proper names will be the same as the code for actual proper names. Reference intuitions about actual proper names are presumably generated by the inference mechanisms, and on the single code hypothesis, the inference mechanisms process pretense representations much as they process belief representations. As a result, the very same mechanisms that produce intuitions about proper names will also produce the intuitions about fictional names. And, since the code is the same, it's no surprise that the intuitions are the same. Our inferences about the relations between Holmes and the associated descriptions in the pretense box will parallel our inferences about the relations between Gödel and the associated descriptions in the belief box. Again, the single code hypothesis does not tell us how to characterize the semantics of fictional names, but it does provide an explanation for why we have the puzzling intuitions we do.³²

iv. Tacit Pretense

The possibility of "dispositional imagining" raises yet another puzzle. When reading *The Great Gatsby*, we "know" that Gatsby couldn't survive decapitation and that he has fewer than 1 billion hairs on his head. But the novel never says such things explicitly and it's unlikely that we

explicitly imagine such things. One reason to think that we have such “implicit” imaginings comes from the fact that we can be surprised when apparent implicit imaginings are violated. When reading *Sir Gawain and the Green Knight* for the first time, it’s plausible that we implicitly imagine that the Green Knight couldn’t survive decapitation, for it comes as rather a shock when he picks up his head after Gawain has chopped it off.

Although it seems that we implicitly imagine that Gatsby couldn’t survive decapitation it is hard to see how to account for this. An account in terms of dispositions seems quite inadequate, as Walton notes: “A person who has recurrent dreams or daydreams of fame and fortune is disposed, between the occurrences, to imagine himself rich and famous. But he need not actually be imagining this nonoccurrently then; he might even imagine the opposite during the intervals.”³³ The fact that John is disposed to imagine himself being famous doesn’t mean that John “dispositionally imagines” being famous.³⁴ Currie goes further and maintains that all imagining is occurrent:

With belief and desire we have the distinction between dispositional and occurrent states. I may believe something I have never thought about and which plays no role in explaining my past or current behavior. Similarly for desire. In these cases we say that the belief or desire is dispositional rather than occurrent; the state is possessed in virtue of the fact that if a relevant situation were to arise you would be likely to behave in a way that would betray the belief or the desire. But there is no comparable category of dispositional imagining, and all imagining is occurrent.... And while I may be disposed to imagine certain things, the possession of such a disposition does not constitute my dispositionally imagining them.³⁵

Currie proceeds to point out that being disposed to run a mile doesn’t mean you are dispositionally running a mile.³⁶ Similarly, then, he suggests, being disposed to imagine that *p* doesn’t mean that you’re dispositionally imagining that *p*. Yet, again, it seems plausible to say that in some sense we implicitly imagine that the Green Knight couldn’t survive decapitation.

This apparent tension about the possibility of “dispositional imagining” is resolved on the single code account. The first thing to note is that Currie’s complaint about “dispositional imagining” finds a parallel in “dispositional believing”. Merely having a tendency to believe that *p* doesn’t entail that one “dispositionally believes that *p*” in the sense we’re considering. For instance, if John has a tendency to believe that he’s fat, it doesn’t follow that he dispositionally believes that he is fat. As a result, the term “dispositional belief” is rather misleading. But the underlying idea seems relatively straightforward. One familiar way to capture the notion is to claim that “dispositional beliefs” are, roughly, beliefs that would be produced immediately by the inference mechanisms under the appropriate circumstances. Lycan uses the term “tacit belief” to capture this sort of notion.³⁷ Lycan’s terminology is useful because it allows us to distinguish *being disposed to believe* from *tacit belief*. John’s disposition to think that he’s fat is not a tacit belief; but his belief that Descartes had fewer than a billion hairs is. Similarly, then, *being disposed to imagine* needs to be distinguished from *tacit pretense*. If tacit beliefs are simply beliefs that would be (under appropriate circumstances) immediately inferred from the current set of beliefs, then tacit pretenses are pretense representations that would be (under appropriate circumstances) immediately inferred from the current set of pretense representations. Since the pretense representations are in the same code as beliefs and are processed by the same inference mechanisms as beliefs, the account of tacit belief is easily appropriated for an account of tacit pretense.³⁸ Given the other things we imagine when reading *The Great Gatsby*, we tacitly imagine that Gatsby has fewer than a billion hairs.

v. *Imaginative resistance*

Richard Moran raises a further, intriguing problem about our emotional and moral responses to fiction, the phenomenon of “imaginative resistance”. While we seem to be willing to imagine the most outrageous things – that pigs can read, that the world is two-dimensional, that Menard tried to rewrite the *Quixote* – we also seem to resist imagining certain things. Moran offers a nice example:

If the story tells us that Duncan was *not* in fact murdered on Macbeth’s orders, then *that* is what we accept and imagine as fictionally true.... However, suppose the facts of the murder remain as they are in fact presented in the play, but it is prescribed in this alternate fiction that this was unfortunate only for having interfered with Macbeth’s sleep that night.³⁹

Although in the former case, we would be willing to imagine that Macbeth did not order Duncan’s murder, in the latter case, Moran suggests, we would not imagine that in the fictional world murdering one’s guest is morally okay. But, if the story tells us that such a murder is morally acceptable, why do we resist imagining this? Part of the puzzle posed by such cases, the part that will be our focus, concerns the psychology of imagination – why do we resist imagining that murder is okay?⁴⁰ Moran indicates that this imaginative resistance is restricted to moral and emotional features, but I’m not so sure. For there are plausibly cases of imaginative resistance in something as amoral and affectless as mathematics. If we are asked to imagine that all of mathematics is just as we think it is, except that $2 \neq 2$, we exhibit imaginative resistance here as well.⁴¹ On the single code hypothesis, this resistance might be explained by the fact that the pretense representation that $2 \neq 2$ would engage our normal inferential systems. And just as our inferential systems would lead us to resist believing that $2 \neq 2$, so too do they lead us to resist imagining that $2 \neq 2$.

If we suppose that pretense representations get processed by our moral response mechanisms, then the single code hypothesis also provides a natural explanation for moral imaginative resistance.⁴² For the moral response mechanisms will treat pretense representations as they treat isomorphic belief representations. In Moran’s case, our moral response mechanisms are engaged by the pretense representations involving the murder of Duncan. *Believing* the facts as presented would prompt a reaction of moral condemnation. Hence, since pretense representations are in the same code, *imagining* the facts activates a moral-condemnation response as well.⁴³ These moral reactions to pretense representations are not pre-empted or eradicated if the play tells us that Macbeth’s actions are morally acceptable, which is just what the single code hypothesis would lead us to expect. For the same holds of our moral reactions to what we believe. If we read a newspaper account of a murder, our moral disapprobation will scarcely be eliminated if the journalist says that the action was morally acceptable.⁴⁴

III. CONCLUSION

I’m sure the basic explanatory strategy afforded by the single code hypothesis is by now painfully evident, so I don’t propose to belabor the matter any longer. But with a little work, I expect that the single code account can also explain several other puzzles about fiction, including puzzles about the attribution of truth in fiction, the assertion of fictional claims, and why fiction seems to admit of “degrees of truth”.⁴⁵ So its potential explanatory sweep is considerable. The single code hypothesis seeks to explain the similarities between believing and imagining by

proposing a strict isomorphism between belief representations and pretense representations. But obviously there are salient and profound discontinuities between imagining and believing, and the single code theorist must somehow accommodate those discontinuities.⁴⁶ That project can't be pursued here, but I hope to have shown that the single code hypothesis does have considerable promise. Indeed, the single code hypothesis seems to provide us with a strikingly powerful account of one corner of the imagination.⁴⁷

SHAUN NICHOLS

Department of Philosophy

College of Charleston

Charleston, SC 29412

Email: nichols@cofc.edu

¹ See, for example, Gregory Currie, "Imagination and Simulation: Aesthetics Meets Cognitive Science," in *Mental Simulation: Evaluations and Applications*, ed. Martin Davies and Tony Stone (Blackwell, 1995); Currie, "The Paradox of Caring," in *Emotion and the Arts*, ed. Mette Hjort and Sue Laver (Oxford University Press, 1997); Robert Gordon & John Barker "Autism and the 'Theory of Mind' Debate," in *Philosophical psychopathology: A book of readings*, ed. George Graham and G. Lynn Stephens (MIT Press, 1994); Paul Harris, *The Work of the Imagination* (Blackwell, 2000), Alan Leslie, "Pretense and Representation: The Origins of 'Theory of Mind'," *Psychological Review* 94 (1987): 412-426; Adam Morton, *Frames of Mind*, (Oxford University Press, 1980); Shaun Nichols and Stephen Stich, "A Cognitive Theory of Pretense," *Cognition* 74 (2000): 115-147.

² The single code hypothesis also has serious problems since there are clear cases in which we respond differently to imagining and believing. In a companion piece to this article, I set out some of the asymmetries that are particularly vexing for the single code hypothesis, and I attempt to provide a way of accommodating the cases without relenting on the claim that imagining and believing are in the same code (Nichols, "Just the Imagination" [forthcoming]).

³ Nichols & Stich, "A Cognitive Theory of Pretense." See also Nichols & Stich, *Mindreading* (Oxford University Press, 2003).

⁴ Alan Leslie, "Pretending and Believing: Issues in the Theory of ToMM," *Cognition* 50 (1994): 211-238.

⁵ See also Currie, "Imagination and Simulation."

⁶ In "A Cognitive Theory of Pretense," Stich and I use the term "Possible World Box" to pick out the pretense box. But the notion of Possible World that we invoke differs from the notion familiar to philosophers. Since this paper is aimed primarily at philosophers, I've adopted a term with less philosophical baggage.

⁷ One possibility is that a large class of the child's beliefs that are consistent with the pretense scenario are simply added to the set of pretense representations. See Nichols & Stich, "A Cognitive Theory of Pretense," pp. 124-25.

⁸ Indeed, as Barbara Von Eckardt has reminded me, it's possible that there are actually multiple kinds of codes in the belief box. That is, some beliefs might be stored in one code, while other beliefs are stored in a different code. Obviously, this possibility complicates the picture that there is a single code for pretense and belief. But the important point is simply that each pretense representation is in the same code as some parallel belief representation. This

qualification does not affect the main points of this paper, so for expository ease, I'll continue to cast the proposal as the "single code" hypothesis.

⁹ This doesn't, of course, mean that the mental processing of pretense representations will be *exactly* parallel to the processing of isomorphic belief representations. After all, the pretense representations do not feed into all the same mechanisms as beliefs. Another obvious difference is that when the inference mechanisms (for instance) receive input from the pretense box, the outputs are typically sent back to the pretense box rather than to the belief box. These qualifications do not, I think, compromise the single code hypothesis, insofar as the qualifications seem independently plausible and well motivated.

¹⁰ Currie, "Imagination and Simulation"; Robert Gordon, "Folk Psychology as Simulation," *Mind & Language*, 1 (1986): 158-170; Alvin Goldman, "Interpretation Psychologized," *Mind & Language* 4 (1989): 161-185; Paul Harris, "From simulation to folk psychology: The case for development," *Mind & Language* 7 (1992): 120-144.

¹¹ Representative defenses of pretend desires can be found in Gordon, "Folk Psychology as Simulation" and Currie and Ian Ravenscroft, *Recreative Imagination* (Oxford: Oxford University Press). Skepticism about pretend desires is defended in Nichols & Stich, "A Cognitive Theory of Pretense" and Nichols, "Review of G. Currie and I. Ravenscroft's *Recreative Minds: Imagination in Philosophy and Psychology*", *Mind* (forthcoming).

¹² Pretend play typically involves overt play behavior that raises additional complications (Nichols & Stich, "A Cognitive Theory of Pretense," section 3.3). But since the focus here will be on imagination without overt behavior, we need not dwell on these complications.

¹³ There is one significant elaboration of the pretense box theory concerning the storage of pretense representations. Most people have a vast knowledge of particular stories. It's plausible that the pretense box contains not only currently developing pretense episodes, but also stores past pretenses. That is, like beliefs, pretense representations can plausibly be stored. For instance, most of us have clusters of Hamlet-representations from our viewings of *Hamlet*. Of course, we don't have these stored as beliefs about a former Prince of Denmark. Rather, the representations that were generated by past viewings of *Hamlet* are likely stored in the pretense box. But clearly some of these stored representations in the pretense box will contradict others. For example, in Steve Erickson's *Tours of the Black Clock*, Hitler lives well past 1945 and grows to be an old man; but in Don DeLillo's *White Noise*, Hitler dies in 1945. As a result, the clusters of stored representations must be kept distinct to avoid inferential catastrophe. Stored pretense representations might be distinguished by indexing, and, with a nod to David Lewis, we can maintain that stored pretense representations can be indexed by worlds. So, for instance, there might be a pretense representation specifying that in the *Tours of the Black Clock* world, Hitler lives to be an old man. There is another pretense representation that says that Hitler did not live to be an old man, but that representation is indexed to the *White Noise*-world.

¹⁴ A number of writers have emphasized this point. See, e.g., Alvin Goldman, "Empathy, Mind, and Morals," *Proceedings and Addresses of the American Philosophical Association*, 66 (1992): 17-41; Goldman, "In Defense of the Simulation Theory," *Mind and Language* 7, (1992): 104-119; Harris, *The Work of the Imagination*; Nichols & Stich, *Mindreading*.

¹⁵ For representative experiments on imagination and fear, see Scott Vrana, Bruce Cuthbert, and Peter Lang, "Processing Fearful and Neutral Sentences: Memory and Heart Rate Change," *Cognition and Emotion* 3 (1989): 179-195; Peter Lang, Daniel Levin, Gregory Miller, Michael Kozak, "Fear Behavior, Fear Imagery, and the Psychophysiology of Emotion: The Problem of

Affective Response Integration,” *Journal of Abnormal Psychology* 92 (1983): 276-306. For reviews of the literature, see Peter Lang, “Cognition and Emotion: Concept and Action,” in *Emotions, Cognition and Behavior*, ed. Carroll Izard, Jerome Kagan, Robert Zajoncs (Cambridge University Press, 1984): 192-226; Harris, *The Work of the Imagination*.

¹⁶ Of course this figure, like the previous figure, leaves out many important functional relations. For instance, the body monitoring system likely plays an important role in emotion generation. The figure is intended to emphasize functional relations that will be particularly important in this paper.

¹⁷ Another open question concerns the possible role of desires in the emotion processing system. The central scientific tradition on emotions, the affect program tradition, does not typically appeal to desires in the explanation of how one goes from a situation to an emotion (see, e.g., Ron Mallon & Stephen Stich “The Odd Couple: The Compatibility of Social Construction and Evolutionary Psychology,” *Philosophy of Science* 67 (2000): 133-154). Rather, the standard explanation of how affect is generated appeals to appraisal processes that seem to take information primarily from the subject’s beliefs about the situation. So for instance, it is appraisal of the details of the situation that produces the affect. Of course, it is possible that desires really play a crucial role here as well, but that the role of desires has not been adequately articulated in contemporary affect theory. If it turns out that desires are important, that will raise additional issues, of course. But it need not pose a problem for the single code hypothesis.

¹⁸ Kendall Walton, *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Harvard University Press, 1990), p. 214.

¹⁹ The literature is vast, but for some of the important discussions, see Noel Carroll, *Beyond Aesthetics: Philosophical Essays* (Cambridge University Press, 2001); Currie, *The Nature of Fiction* (Cambridge University Press, 1990); Currie, *Image and Mind* (Cambridge University Press, 1995); Susan Feagin, *Reading with Feeling: The Aesthetics of Appreciation* (Cornell University Press, 1996); Tamar Gendler, “The Puzzle of Imaginative Resistance,” *Journal of Philosophy* 97 (2000): 55-81; Hjort and Laver, *Emotion and the Arts*; Peter Lamarque and Stein Olsen, *Truth, Fiction, and Literature* (Oxford University Press, 1994); Jerrold Levinson, *The Pleasures of Aesthetics* (Cornell University Press, 1996); David Lewis, “Truth in Fiction,” in his *Philosophical Papers, vol. I* (New York: Oxford University Press, 1983); Richard Moran, “The Expression of Feeling in the Imagination,” *Philosophical Review* 103 (1994): 75-106; Kendall Walton, *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Harvard University Press, 1990); Walton, “Spelunking, Simulation and Slime: On Being Moved by Fiction,” in *Emotion and the Arts*.

²⁰ This formulation of the puzzle follows Currie, “The Paradox of Fiction,” p. 65.

²¹ In “The Paradox of Fiction,” Currie offers an off-line simulation explanation of the puzzle of fiction and emotions, and since I take off-line simulation theory to be a kind of single code theory, Currie’s treatment counts as a single-code account. Aaron Meskin and Jonathan Weinberg provide an alternative to Currie’s simulation-approach to the problem that is very much in the spirit of the account presented here. See their “Emotions, Fiction, and Cognitive Architecture,” *The British Journal of Aesthetics* (2003).

²² Walton, *Mimesis*, p. 284.

²³ It’s not entirely clear what the mechanisms are subserving this effect in the case of belief. But the single code theorist need only suggest that, the mechanisms that produce this effect for belief are also implicated by pretense representations.

-
- ²⁴ For further discussion of iteration puzzles, see Nichols, “Imagination and the Puzzles of Iteration,” *Analysis* 63 (2003).
- ²⁵ Currie, *The Nature of Fiction*, p. 151.
- ²⁶ Saul Kripke, *Naming and Necessity* (Harvard University Press, 1980), p. 81.
- ²⁷ Kripke, *Naming and Necessity*, p. 157. Kripke’s point here influences David Lewis’ account of truth in fiction – see Lewis, “Truth in Fiction,” pp. 265, 276.
- ²⁸ Kripke, *Naming and Necessity*, pp. 83-84.
- ²⁹ Currie, *The Nature of Fiction*, p. 165.
- ³⁰ Kripke, *Naming and Necessity*, p. 61.
- ³¹ These sorts of views are suggested in Walton, *Mimesis* and in unpublished work by Kripke.
- ³² Whether this explanation of the intuitions impacts the semantic project is a separate question that I will not consider here. For one account of the semantics of fictional names that dovetails with the account of the imagination presented here see Stacie Friend, “The Great Beetle Debate: A Study in Imagining with Names” (forthcoming).
- ³³ Walton, *Mimesis*, p. 17.
- ³⁴ Walton does maintain that there are “non-occurrent” imaginings, and merely raises the above as a problem for a simple dispositional analysis of non-occurrent imaginings (Walton, *Mimesis*, 16-18).
- ³⁵ Currie, “Imagination and Simulation,” p. 156.
- ³⁶ *Ibid.*
- ³⁷ William Lycan, *Judgement and Justification* (Cambridge University Press, 1988). See also Jerry Fodor, *Psychosemantics* (MIT Press, 1987), p. 22.
- ³⁸ Since my focus is not tacit belief, I don’t really want to commit to one particular account of tacit belief. Rather, the suggestion is that *whatever* the right story is about tacit belief, there should be an analogous story about tacit pretense.
- ³⁹ Moran, “The Expression of Feeling in the Imagination,” p. 95.
- ⁴⁰ But this is not the only puzzle raised over such cases. Indeed, Moran couches the problem largely in terms of what we accept as true in the fiction, and Walton has argued that the issues concerning truth in fiction need to be distinguished from the question about the psychology of the imagination. Walton’s own discussion of these matters focuses primarily on our attitudes concerning what is true in the fiction (“Morals in Fiction and Fictional Morality,” *Aristotelian Society Supplementary Volume* 68 (1994), 27-50. See also *Mimesis*, pp. 154-5).
- ⁴¹ Cf. Edward Craig, “The Problem of Necessary Truth,” in *Meaning, Reference, and Necessity*, ed. Simon Blackburn (Cambridge University Press, 1975).
- ⁴² We must also suppose that our moral judgment mechanisms differ from our etiquette judgment mechanisms, for we typically exhibit no resistance to imagining alien etiquette (cf. Hume, “On the Standard of Taste”). But it’s independently plausible that the mechanisms differ, and one explanation for the difference appeals to the role of emotional responses in moral judgment (e.g., Nichols, *Sentimental Rules* [Oxford University Press, forthcoming]).
- ⁴³ However, the phenomena surrounding imaginative resistance are a complex lot. For instance, in some cases, we find *imaginative compliance*. For example, we seem to be able to imagine things that are immoral in cases like the following:
- If it were okay for Macbeth to order Duncan’s murder, would Lady Macbeth’s machinations also be okay?

These kinds of cases might count as a *prima facie* difficulty for the single code hypothesis (see Nichols, “Just the Imagination”).

⁴⁴ In her recent discussion of this problem, Tamar Gendler tells a rather different story. She maintains that we exhibit imaginative resistance because we don’t *want* to imagine morally objectionable perspectives. She writes, “the primary source of imaginative resistance is not our *inability* to imagine morally deviant situations, but our *unwillingness* to do so” (Gendler, “The Puzzle of Imaginative Resistance,” p. 56). Gendler provides compelling examples in which we do indeed resist imagining certain things because we don’t want to (see e.g. pp. 73-74). But the single code hypothesis suggests that there is also a more basic *architectural* explanation for much imaginative resistance. The inferential and emotional devices that lead us to resist certain beliefs are also triggered by pretense representations. (Derek Matravers offers an account of imaginative resistance that resembles the single code explanation in some ways [see his “Fictional Assent and the (so-called) ‘Puzzle of Imaginative Resistance,’” in *Imagination, Philosophy, and the Arts*, ed. Matthew Kieran and Dominic Lopes (Routledge, 2003)]. However, to explain the phenomena, Matravers invokes a fictional narrator and maintains that resistance occurs because we judge that the fictional narrator lacks authority in the target domain. The single code account need not advert to these controversial resources to explain imaginative resistance.)

⁴⁵ For presentations of these puzzles see Currie, *The Nature of Fiction*, pp. 53ff., Walton, *Mimesis*, pp. 207-8, and Currie, *The Nature of Fiction*, p. 90.

⁴⁶ For a broad taxonomy of symmetries and asymmetries in imagination and belief see Tamar Gendler, “Quarantining and Contagion, Mirroring and Disparity: On the Relation between Pretense and Belief,” in *Imagination, Philosophy, and the Arts*, ed. Matthew Kieran and Dominic Lopes (Routledge, 2003).

⁴⁷ This paper was presented at the 2002 NEH Institute on Art, Mind, and Cognitive Science. I’d like to thank the audience for useful comments. I’m also grateful to Greg Currie, Stacie Friend, Tamar Gendler, James Harold, Aaron Meskin, Adam Morton, Jenefer Robinson, Barbara Von Eckardt, Ken Walton, Jonathan Weinberg, an anonymous referee, and especially Steve Stich for discussion and comments on earlier drafts of this paper.