

Forthcoming in *Mind & Language*

Just the Imagination: Why Imagining Doesn't Behave Like Believing^{*}

SHAUN NICHOLS

Abstract: According to recent accounts of the imagination, mental 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. That is, such a mechanism should produce similar outputs whether its input is the belief that p or the pretense representation that p . Unfortunately, there seem to be clear counterexamples to this hypothesis, for in many cases, imagining that p and believing that p have quite different psychological consequences. This paper sets out some central problem cases and argues that the cases might be accommodated by adverting to the role of desires concerning real and imaginary situations.

^{*} For helpful comments and discussion, I'd like to thank Jonathan Cohen, Greg Currie, Stacie Friend, Tamar Gendler, Aaron Meskin, Steve Stich, and two anonymous referees for *Mind & Language*.

Address for correspondence: Department of Philosophy, University of Utah, Salt Lake City, UT, 84112, USA.

Email: snichols@philosophy.utah.edu

Recent cognitivist accounts of the imagination largely converge on the idea that there is a ‘single code’ for both imagining and believing. According to the single code hypothesis, psychological 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. So, a psychological mechanism that takes input from the imagination and from belief should generate importantly similar outputs for the pretense representation that *p* and the belief that *p*. This view of the imagination is shared by both simulation theorists (Gordon & Barker 1994; Currie 1995; Harris 2000) and their foes (Leslie 1987; Nichols & Stich 2000).

The single code hypothesis provides a unified explanation for a wide range of puzzles concerning fiction and the imagination. The best known result is that it helps solve the paradox of fiction (Currie 1997; Meskin and Weinberg 2003); but the single code hypothesis can also be recruited to explain imaginative resistance, puzzles about iterated imagination, intuitions about fictional names and a number of other established problems about fiction and imagination (Nichols 2003, 2004a, forthcoming). It is, perhaps, the most productive idea about the propositional imagination that anyone has ever had. Unfortunately, the single code hypothesis also faces some serious obstacles. The goal of this paper is to set out the problems and consider the available solutions. In the first section of this paper, I will describe the single code hypothesis a bit more fully. In section 2, I will set out the problem cases. Then in the final section, I’ll consider various options for rescuing the single code theory from the problems.

1. The Single Code Hypothesis

To explain the single code hypothesis, it's important to get clear about some background assumptions concerning the cognitive architecture of the imagination. First, I will adopt the representationalist approach that is common in this area and say that propositional imagining involves 'pretense representations'. To imagine that Hamlet is a prince is to have a pretense representation with the content *Hamlet is a prince*.¹ Second, I will assume that a pretense representation can have exactly the same content as a belief. Indeed, on contemporary accounts of the imagination, one can simultaneously have a belief with the content *p* and a pretense representation with the content *p*. This is nicely illustrated in an experiment of Alan Leslie's. Leslie had young children watch as he pretended to pour tea into two (empty) cups. Then he picked up one of the cups, turned it over and shook it, turned it back right side up and placed it next to the other cup. The children were then asked to point at the 'full cup' and at the 'empty cup'. Both cups were really empty throughout the entire procedure, but two-year-olds reliably indicated that the 'empty cup' was the one that had been turned upside down and the 'full cup' was the other one (Leslie 1994). On the most natural interpretation of this, the child is *imagining that the cup is empty*. But the child also, of course, *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. Rather, contemporary accounts of the imagination maintain 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 (see also Currie 1995, 1998). In our account of pretense and

¹ Those who eschew mental representations entirely, e.g., behaviorists, would reject this claim outright. The cognitivist accounts of the imagination to be considered here are unabashedly steeped in mental representations.

imagination, Stephen Stich and I exploit the familiar illustrative device of using boxes to represent functional groupings, and we propose that there is, in addition to a belief box and a desire box, a ‘pretense box’ (Nichols & Stich 2000). The final important background assumption is that pretense representations interact with some of the same psychological mechanisms that belief representations interact with. Consider again Leslie’s experiment. Virtually all of the children in Leslie’s experiment responded the same way when asked to point to the ‘empty cup’. How are these orderly patterns to be explained? The prevailing cognitivist view is that the pretense representations are processed by the same inference mechanisms that operate over real beliefs.²

With the functional architecture in place, we can now articulate the core feature of cognitivist accounts 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 from Leslie (1987, 417). Of course, no one knows 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.³ But the critical point can be made without

² Of course, there are probably also mental mechanisms that take input from the belief box but not from the pretense box. For instance, at least on the story that Stich and I tell, the practical reasoning mechanism takes beliefs and desires as inputs, but it does not take pretense representations as input (Nichols & Stich 2003).

³ 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

giving further detail about what the code is. The key point is just that, 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.⁴ 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. As a result, while off-line simulation theorists don’t frame the issue

pretense representation is in the same code as some parallel belief representation. This qualification does not affect the main points of this paper, however, so for expository ease, I’ll continue to cast the proposal as the ‘single code’ hypothesis.

⁴ 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. Introspection presents yet another kind of case. For when I introspect my pretense representations, the output is a belief about what I’m pretending, and when I introspect my beliefs, the output is a belief about what I believe. So, if a single introspection mechanism processes both pretense representations and beliefs, then this mechanism treats pretense representations and beliefs in importantly different ways. These qualifications do not, I think, compromise the single code hypothesis, insofar as the qualifications seem independently plausible and well motivated.

quite this way, they are plausibly viewed as single code theorists. For they maintain that various mental mechanisms process 'pretend beliefs' much like real beliefs (e.g., Currie 1995, Gordon 1986, Goldman 1989, Harris 1991). 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 (e.g. Gordon 1986; Currie & Ravenscroft 2002). Those additional stipulations are consistent with, but not required by the single code hypothesis. There are, I think, good reasons to be skeptical of pretend desires (Nichols & Stich 2000; Nichols 2004b), so here I will restrict the single code hypothesis to the narrower hypothesis.

On the single code hypothesis, then, 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 is a general hypothesis about psychological mechanisms, not specific to the inference mechanisms. If, for instance, emotion systems receive input from the imagination, the single code hypothesis predicts that these systems should produce affective output similar to the output that would be produced by isomorphic beliefs. So, if the pretense representation that *p* gets processed by an affective mechanism, the affective outputs should parallel those that would follow from the belief that *p* (see, e.g., Goldman 1992a, 1992b; Harris 2000; Nichols & Stich 2003). This claim gains support from a wide range of empirical findings (see Lang 1984, Harris 2000 for reviews). For instance, in one experiment, subjects were presented with an imaginary scenario involving a snake encounter. 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 (Lang et al. 1983). So the particular character of the subject's affective response system seems to be reflected in their affective responses to imagination.⁵

As noted in the beginning, the single code hypothesis helps explain both the paradox of fiction and imaginative resistance. It will be important for what follows to say just a bit more about this here. The paradox of fiction might be put as an inconsistent triad:

I pity Lear.

To pity something, you have to believe it exists.

I don't believe Lear exists.⁶

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. Single code theories promise a solution to at least the psychological puzzle of how fiction can cause affective response (Currie 1997; Meskin & Weinberg 2003). Affective systems can receive

⁵ It's important to note that the single code theorist can be neutral with respect to whether the physiological arousal counts as genuine fear or merely quasi-fear (see Currie 1997). The point is simply that the output of the fear mechanism is importantly similar when the input is the belief that *p* and when it is the pretense representation that *p*. In some contexts, this qualification is important; here, however, it isn't. So for ease of exposition, I will henceforth omit the qualification that the affective mechanisms might only produce 'quasi-emotions' in response to input from the imagination.

⁶ This formulation of the puzzle follows Currie (1997, p. 65).

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 inputs 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 provides a natural solution for the problem 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. Richard Moran sets the problem nicely:

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 (Moran 1994, p. 95).

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 occupy us, concerns the psychology of imagination – why do we resist imagining that murder is okay? Consider first a mathematical example.⁷ If we are asked to imagine that all

⁷ Moran indicates that imaginative resistance is restricted to moral and emotional features, but it seems that a similar phenomenon occurs with mathematical examples, as elaborated above.

of mathematics is just as we think it is, except that $2 \neq 2$, we exhibit imaginative resistance here as well (cf. Craig 1975). On the single code hypothesis, this resistance might be explained by the fact that the occurrent pretense representation that $2 \neq 2$ would engage our normal inferential systems. And just as our inferential systems would lead us to reject the occurrent belief that $2 \neq 2$, so too do they lead us to reject the occurrent pretense representation that $2 \neq 2$. Similarly, then, if we suppose that pretense representations get processed by our moral response mechanisms, the single code hypothesis has a natural solution to the problem of moral imaginative resistance. For the moral response mechanisms will treat pretense representations as they treat isomorphic belief representations. Having the occurrent *belief* that Macbeth ordered Duncan's murder leads us to regard the action as immoral. Hence, since imagining is in the same code, having the occurrent pretense representation that Macbeth ordered Duncan's murder leads us to regard the action as immoral in the imagination. That would explain why we resist alien moral judgments even in imagination.

2. Asymmetries

Despite all the systematic similarities between imagining and believing, there are also obvious cases of systematic discrepancies.⁸ Of course, the single code hypothesis does not claim that imagining that p and believing that p *always* have the same psychological effects. Some discontinuities between imagination and belief aren't troubling in the least. For instance,

⁸ See Gendler (2003) for a broad taxonomy of symmetries and asymmetries in imagination and belief.

imagining that my dog died and believing that my dog died will have significantly different consequences. For one thing, the affect generated by the belief that my dog died will presumably endure longer than the affect generated by my imagining that my dog died. For one can simply stop imagining the dog's demise, but this option is less readily available for belief. The single code hypothesis can also accommodate the fact that our affective reactions to fiction are often less profound than those to the real world. For instance, my response to the death of Desdemona is much weaker than the response I would have to the death of my sister. But the single code theorist can point to important differences between what I imagine concerning Desdemona and what I would believe if my sister died. One crucial difference is in identification. Even in the imagination, I don't identify with Desdemona the way I identify with my sister, and it's plausible that this would contribute to differences in affective responses. More tellingly, the emotional responses I have to *imagining* that my sister has died are more profound than the emotional responses I have to imagining that Desdemona has died. As a result, these kinds of cases aren't worrying because we can explain the discontinuities without yielding on the single code hypothesis. However, there are some kinds of asymmetries that are much harder to explain.

The first class of cases involves what we might call *discrepant affect*. Sometimes our affective responses to imaginings differ in striking ways from our responses to belief. The clearest cases come from black comedy. At the end of *Dr. Strangelove*, we imagine that all human life is about to be destroyed, and we find this amusing in the context of the film. Presumably this is not how we would react if we had the real belief that all human life is about to be destroyed. Perhaps if we really believed that all human life was about to be destroyed, we could find some humor in the situation, but surely this would not be the predominant emotional

response.

In the other class of cases that I'll worry over, the discontinuity is that imagining that p sometimes produces *no affect* when believing that p would generate considerable affect. We can look close to home for a wealth of such examples – philosophical thought experiments. Philosophical thought experiments are often unaccompanied by the affective consequences that would result if we really believed the situation described in the thought experiment. Consider, for example, the following thought experiment:

Imagine that you're red-green colorblind and that all sentient life in the universe except for you is destroyed. In that case, does the color red still exist?

If I really came to believe that all other sentient life in the universe was destroyed, I wouldn't trifle with questions about secondary qualities, and the affective consequences would be profound. In my experience of using these types of thought experiments in class, I've never seen any student react with the emotional responses that would be expected if the student really had these beliefs. And I doubt that my students are unusually callous.

What these problems share is that in each case it seems the most plausible thing to say is that our affective mechanisms treat the pretense representation that p systematically differently from the belief that p . And that's exactly what the single code hypothesis says won't happen.

3. Solutions

The cases of discrepant affect and absent affect present a serious challenge to the single code hypothesis, and I don't pretend to have a definitive solution. In this section, I'll consider two unsatisfying responses to the asymmetries and I'll sketch a third response that strikes me as

promising.

3.1 Bifurcating the Imagination

One natural way to preserve the single code hypothesis while addressing the problem cases is to maintain that there are different *kinds* of imagination. That is, there are two different propositional imagination systems, and the single code hypothesis only applies to one of these types of the imagination.⁹ There is an important recent precedent for the view that there are two types of the imagination. To explain the precedent, we need to return to the issue of imaginative resistance. For resistance is fickle. As Moran notes in his discussion of imaginative resistance, although we resist imagining the immoral in some cases, in other cases, we show no such resistance. Consider the following:

If it were okay for Macbeth to order Duncan's murder, would Lady Macbeth's machinations also be okay?

This seems to be a case of *imaginative compliance*, for here we seem capable of imagining something that we wouldn't believe. To solve the problem of imaginative compliance, Moran distinguishes between hypothetical imagining and dramatic imagining (1994, 104). It's not entirely clear whether this is supposed to be a distinction between different kinds of cognitive activities, but this is suggested by Moran's further discussion. The kind of activity involved in hypothetical imagining is characterized as follows:

⁹ Obviously one needs to distinguish the kind of imagining that I've focused on (propositional imagining) from visual imagining and the like. But that distinction won't help with the current problems.

Hypothetical reasoning involves seeing what would follow from the truth of some proposition. It does not involve either feigning belief in that proposition or determining what would follow from the fact of one's believing it. There need be no reference to oneself, either as believer or as any sort of psychological subject, and one does not determine the truth of a counterfactual by imagining 'what it would be like' to believe the antecedent... in the case of ordinary counterfactual reasoning... one only needs to be provided with the proposition in order to reason from the assumption of its truth (105).

Dramatic imagining is characterized much differently:

imagination with respect to the cruel, the embarrassing, or the arousing involves something more like a point of view, a total perspective on the situation, rather than just the truth of a specifiable proposition. And imagining along these lines involves something more like genuine rehearsal, 'trying on' the point of view, trying to determine what it is like to inhabit it (105).¹⁰

Moran's distinction between hypothetical and dramatic imagining is intriguing, and if it explains imaginative compliance, it would also presumably explain the absent affect cases. For we could say that cases of absent affect and imaginative compliance involve only the hypothetical imagination. Affective response and imaginative resistance occur only with the deployment of the dramatic imagination. But I don't think that bifurcating the imagination will fully address the problems we face. First, it's not sufficient merely to claim that there are two kinds of imagining, only one of which involves the deep engagement we see reflected in imaginative resistance and affective response. We already knew that sometimes our imagination

¹⁰ Tamar Gendler makes a similar distinction between supposition and imagining to address cases of imaginative compliance (Gendler 2000, 80-1).

engages our moral/emotional responses and sometimes it doesn't. What we want is an independent characterization of when and why the deeper engagement occurs. Merely labeling the deeper engagement *dramatic* doesn't advance our understanding. In order for the bifurcation move to address our problem, we need a substantive story about the psychological nature of the distinction. Moran's discussion is brief, but it suggests a couple of different ways one might draw the distinction.

One possibility is that the hypothetical imagination is distinguished from the dramatic by the domains to which Moran appeals – 'the cruel, the embarrassing, or the arousing'. But it is clear that appealing to these domains will not suffice for drawing the distinction. For there will be cases of imaginative compliance that do involve the cruel. For instance, we seem to be able to imagine that it's okay to torture innocent people in cases like the following: 'If it were okay to torture innocent people, would it be okay to torture guilty people?' Similarly for the embarrassing: 'Would it be more embarrassing to be naked in front of 100 people or 1000 people?' So it doesn't seem to be the *domain* (cruel, embarrassing, etc.) that divides 'hypothetical' from 'dramatic' imagining.

Another way to draw the distinction is to claim that dramatic imagining demands that we *deliberately* engage the content deeply. Perhaps this is what Moran is suggesting when he says that dramatic imagining might involve "something more like genuine rehearsal, 'trying on' the point of view, trying to determine what it is like to inhabit it" (105). Again, it's tempting to think this will explain imaginative compliance and the cases of absent affect. We might maintain that when we are asked to carry out some hypothetical reasoning, as in the secondary qualities case, we do not attempt the dramatic rehearsal, and this is why the affect is absent. Unfortunately, this approach to the hypothetical/dramatic distinction still fails to solve our

problems. Try processing the following only in ‘hypothetical mode’:

Would the tablecloth be dirty if the diner had eviscerated a cat at the table and eaten its dripping entrails?

Some disgust response to this question is largely inevitable. Indeed, I feel obligated to apologize for subjecting the reader to the sentence. The disgust response is typical here even if we do not deliberately try to engage the content deeply. To solve the hypothetical task doesn’t even require deep engagement with the content. Certainly we can solve the hypothetical task here much more easily than we can in some cases that do not generate affective response, like trying to figure out whether *red* would exist if I were colorblind and alone in the universe. So, without some other account of the distinction between hypothetical and dramatic imagining, we still lack an explanation for the absent affect cases.

Finally, even if the distinction between hypothetical and dramatic imagining could solve the problems of absent affect and imaginative compliance, it wouldn’t explain cases of *discrepant* affect. So, when we are amused by black comedies, our affect doesn’t fit the situation. Saying that we have two different kinds of imagination does not explain why one type produces affect that doesn’t fit the situation.¹¹

¹¹ The point of the foregoing has been to show that the hypothetical/dramatic distinction will not solve our problems. But some such distinction among imaginative capacities seems to be required to explain our psychological relationships with contradictions. For it’s plausible that while we can’t imagine explicit contradictions, we can take explicit contradictions as assumptions when doing formal logic. One possibility is that taking assumptions in formal logic is a learned skill that bypasses the natural imaginative capacities (cf. Gendler 2000 on ‘supposing’). People unskilled in logic don’t know *how* to reason from contradictions, even if

Although the hypothetical/dramatic distinction fails to explain the difference between the disgusting tablecloth case and the bland secondary quality case, the two cases do bring out one important factor in the generation of affective response. In a lament on sexual jealousy, Elvis Costello wails that it's the 'stupid details' of his lover's sexual liaisons that torment him. Indeed, the stupid details seem to be exactly what prompts disgust in the tablecloth case and such disturbing details are precisely absent from the secondary qualities case. Of course, even if the important difference between the cases comes down to the representational details, this really only pushes our problem back a step. For the problem is that in the secondary qualities case, being presented with exactly the same details produces different affect depending on whether one imagines or believes that all sentient life has been destroyed. That's the challenge for the single code hypothesis that we have yet to answer.

3.2 Associationism

Rather than adopt a multiple-imagination approach, one might eschew the single code hypothesis altogether and adopt the view that the imagination generates associations and that these associations produce the emotional responses (cf. Nichols et al. 1996, 61-2). Since associations are notoriously flexible, it seems quite possible that imagining that *p* would produce different associations than believing that *p*. This might then explain why imagining that *p* and believing that *p* can produce different affective response. The representations that trigger the emotions are their natural imaginative endowment is fully functioning. The crucial point for present purposes, though, is that such a distinction between assuming and imagining still won't address the asymmetries that plague the single code hypothesis. For we can't assimilate all the problematic cases of discrepant affect and absent affect to assumption taking in formal logic.

produced by associations, and imagining that p produces different associations from believing that p . This approach might capture all of our problem cases – in each case, imagining that p sets up different associations than believing that p , hence it's no surprise that the emotional responses are not the same.

The pure associationist line clearly can dispense with our problem cases, but insofar as it rejects the single code hypothesis, the associationist account sacrifices the rich explanatory power of the single code theory. So, the pure associationist account does not provide a very satisfying story. For it fails to give a motivated explanation for the wide swath of cases in which the effects of imagining that p closely track the effects of believing that p .

There is an obvious compromise looming. Since the associationist account explains the problem cases and the single code hypothesis explains the non-problem cases, we should, perhaps, adopt a mixed theory. The single code hypothesis is perfectly right, but associations often distort the processing, as in the problem cases. However, this compromise seems rather *ad hoc*. Further, the simple appeal to associative processes tells us too little about when imagining that p will have different associations from believing that p . In what follows, I want to try to give a fuller explanation of the asymmetries by appealing to the role of desire.¹²

3.3 Differential effects of desires

There are, of course, important differences in the cognitive effects of beliefs and imaginings on the single code account. For example, on the theory that Stich and I defend, believing that p plays a central role in arriving at decisions and instrumental desires (Nichols & Stich 2000).

¹² Depending on one's views about what counts as an association, the account that I give might be regarded as one elaborated version of the associationist account.

Imagining that p does not play a parallel role. Further, we have vast control over what gets *into* the imagination. Not so with belief. A full explanation of the different effects of imagining that p and believing that p will certainly advert to these important differences. But I suspect that the obstacles facing the single code account cannot be addressed adequately by appeal simply to these differences. There is another resource available, however, that does seem promising: desires. Ultimately, I'll suggest that we can explain the problem cases by appealing to the fact that we have different desires about the real situations and the imaginary situations, and these different desires will influence the inferences and memories that are elicited. But to make the case for this, I first need to review how desires affect inference and recall.

When we acquire a new belief, the course of drawing inferences is a function of our prevailing desires and interests. For a crude illustration, consider two zealous sports fans, pulling for different teams A & B in a double elimination tournament. The two fans have all the same beliefs about the tournament. But on learning that both A & B lost in the second round, the A-fan is more likely than the B-fan to draw immediately the inferences for A's future match-ups in the tournament. Further, what is recalled and the order of recall is plausibly affected by desires. If the A-fan and the B-fan learn that A & B are both eliminated in the penultimate round, this will likely elicit different memories in the different fans. The A-fan's memories will involve team A more centrally than the B-fan's memories.

So, when we acquire a new belief, *what we care about* structures both the inferences and the memory activations that ensue. I've used crude examples to make this point. But there is a large and sophisticated body of work in social psychology that charts how motivation influences both inference and recall (see Kunda 1999 for an excellent review). For instance, in one experiment, subjects were induced to think either that being introverted was desirable or that

being extraverted was desirable. Then subjects were asked to recount memories of past behaviors that indicated the extent to which they were introverted or extraverted. Subjects induced to find introversion desirable were more likely to recount introverted memories first and to produce more introverted memories than subjects induced to find extraversion desirable (Sanitoso, Kunda & Fong 1990). Empirical work has also indicated that people with different motives will draw different inferences from the same information. In one experiment, two groups of subjects were given information indicating either that people in their group were more prosocial (desirable) or that they were less prosocial (undesirable) than a different group. The information presented to all of these subjects was based on a small sample. However, the subjects who were presented with an undesirable conclusion (that their group was less prosocial) were significantly more likely to call into question the adequacy of the sample size than the subjects who were presented with the desirable conclusion (Doosje et al 1995).

A rather different tradition of work points in a similar direction. Research on text processing has explored the kinds of spontaneous inferences subjects make during reading (see e.g., Goldman et al. 1999). One strand of this large literature indicates that a reader's goals has significant effects on the inferences generated (e.g., Noordman et al. 1992; Zwaan et al 1995). For instance, in one experiment, participants read adapted *Scientific American* articles under two conditions – either with a goal of entertainment or with a goal of studying for an exam (van den Broek et al. 2001). Subjects were trained to verbalize their thoughts while reading (see Ericsson & Simon 1993). Readers with a study goal tended to produce more explanatory and predictive inferences (e.g. 'Oh, this is related to what I read earlier about how magnetic field affects wave direction'; 'I bet the turtles will use sense of smell to return to their nesting place'), and readers with an entertainment goal tended to produce more evaluations (e.g., 'These turtles are really

intelligent creatures’) (van den Broek et al. 2001, 1084-5). For our purposes, this evidence serves to confirm the commonsensical idea that what we want affects the pattern of our spontaneous inferences.

All of the above cases plausibly involve desires about the real world. As note earlier, some theorists maintain that, in addition to real desires about the real world, we also have pretend desires or ‘desire-like imaginings’ (e.g., Currie & Ravenscroft 2002). If there are such pretend desires, that would provide the single code theorists with an important resource for addressing the problem cases. Since I am skeptical of the existence of pretend desires, I cannot avail myself of their assistance. Rather, I need to appeal strictly to real desires. Fortunately, it’s plausible that we have real desires about imaginary situations. For instance, we have desires about what happens to characters in a fiction. We want the hero to prevail, the villain to suffer defeat. We also have broader desires about various fictions – e.g. some like stories with Hollywood endings, others find such stories tiresome. And we have more complicated desires about the conformity of a narrative to a genre – we might want a tragedy to sustain a tragic narrative. Indeed, sometimes our real desires about fiction can even be in tension. When watching *Othello*, I both want it to be the case (fictionally, of course) that Othello not kill Desdemona, and I also want it to be the case that the narrative be tragic.¹³

The last several paragraphs recount uncontroversial features about our real desires. But these familiar features of desires might provide a basis for rescuing the single code theory from the problem cases. Just as two individuals often have different desires about what happens in the real world, a single individual’s desires about what happens in the world often differ from her

¹³ Currie and Ravenscroft appeal to desire-like imaginings to explain this tension (2002, 21-2).

But again, I’m trying to get by without adverting to those resources.

desires about what happens in an imaginary scenario. I have strong and persistent desires for the health of my family, my friends, my colleagues. However, I often have no closely parallel desires about what happens in an imaginary scenario. Since what we desire affects recall and inference, this will have differential effects on what gets elaborated in the imagination. This point might be illustrated by considering the inferences and memories that would be elicited under two different conditions:

1. Someone walks into your office and says: Imagine that everyone outside of this room is dead, what would Utilitarianism say about the importance of our interests?
2. Someone walks into your office and says: Everyone outside of this room is dead; what does Utilitarianism say about the importance of our interests?

In the latter case, if you believed the messenger, you would inevitably process the implications about your loved ones and you would likely recall particular things about them. But in the former case, such inferences and recollections are far from inevitable. These differences in inference and recall might be explained by the fact that we have different desires concerning the imaginary situation and the real situation. In both (1) and (2) we are encouraged to have the desire to figure out the entailments of Utilitarianism. In case (2), our desires about the real world would swamp any desire we have to work out the Utilitarian calculus. Hence, our inferences and recollections would be guided by these prevailing desires. In case (1), by contrast, we plausibly have no particularly pressing desires about the inhabitants of the imaginary scenario. As a result we are not compelled to draw the inferences and recollections that would follow in case (2). Rather, our desire to answer the question about Utilitarianism can be pursued without the intrusion of salient desires and concerns about the inhabitants of the imaginary situation.

As a result, there might be a great disparity between the representations that would ensue

if we came to believe that p and the representations that would ensue if we imagined that p . Imaginings can be constrained, filtered and directed in all sorts of ways that are not available to beliefs. For our desires about the imaginary scenario will depend on the context, the intent of the author, the tone of the work, the point of the thought experiment, and so on. Our desires about the real world are much less flexible.

With the inferential and mnemonic influences of desires in hand, we can finally return to our problem cases. Our desires about what happens in the real world guide the way our beliefs get elaborated on acquiring new beliefs; similarly, our desires about what happens in an imaginary scenario guide the way pretense representations get elaborated. This difference might then explain the problematic asymmetries between imagining and believing. The explanation for the asymmetries is *not* that the affective mechanism itself responds differently to imagining that p and believing that p . Rather, the asymmetries arise because the affective mechanism is sent quite different *input* depending on whether one imagines that p or believes that p . For instance, in the cases of absent affect (e.g., the secondary qualities case), what we are imagining is extremely sparse. We want primarily to solve the hypothetical problem and accordingly focus on the minimal conditions. In those cases, then, the disturbing inferences are not drawn out and hence they do not reach the affective system. Our desires about the real world are obviously much richer and less accommodating. So if we actually believed that everyone was killed, our desires would, as noted above, inevitably lead us to draw out the consequences for our loved ones. And then those inferential consequences will trigger the affective system.

The discrepant affect cases might get a parallel explanation. When watching *Dr. Strangelove*, the set of inferences and activations we have about the imaginary scenario is shaped by our desires about what happens in that imaginary scenario. When it comes to the real world,

we have powerful and consuming desires for the survival of human life. Those desires about the real world infuse our lives and, if we were led to believe that human life is about to end, the desires would compel us to draw out the ghastly inferences. When it comes to black-comedy, we typically do not have such powerful desires for the preservation of human life in the imaginary scenario. Hence, we are not compelled to draw out disturbing inferences like *billions of innocent people will die horrifically painful deaths*. Rather, genre considerations make us want to focus instead on Slim Pickens' exuberant missile ride.

Differential effects of desires about real and imaginary situations almost certainly contribute to some of the striking asymmetries between imagining and believing. It would be premature to maintain that the role of desires provides a full explanation of the cases that trouble the single code hypothesis. But the availability of a desire-based explanation of the asymmetries can, I hope, help to preserve confidence in the idea that imagining and believing do share a common code.

Department of Philosophy

University of Utah

References

Craig, E. 1975: The problem of necessary truth. In *Meaning, Reference, and Necessity*, S.

Blackburn (ed.). Cambridge: Cambridge University Press.

Currie, G. 1995: Imagination and simulation: Aesthetics meets cognitive science. In *Mental*

Simulation: Evaluations and Applications, A. Stone and M. Davies (eds.). Oxford: Basil

Blackwell.

- Currie, G. 1997: The paradox of caring. In *Emotion and the Arts*, M. Hjort and S. Laver (eds.). Oxford: Oxford University Press.
- Currie, G. 1998: Pretence, pretending and metarepresenting. *Mind & Language*, 13, 35-55.
- Currie, G. and Ravenscroft, I. 2002: *Recreative Imagination*. Oxford: Oxford University Press.
- Doosje, B., Spears, R. & Koomen, W. 1995: When bad isn't all bad: Strategic use of sample information in generalization and stereotyping. *Journal of Personality and Social Psychology*, 69, 642-655.
- Ericsson, K. & Simon, H. 1993: *Protocol Analysis: Verbal Reports as Aata*. Cambridge, MA: MIT Press.
- Gendler, T. 2000: The puzzle of imaginative resistance. *Journal of Philosophy*, 97, 55-81.
- Gendler, T. 2003: Quarantining and contagion, mirroring and disparity: On the relation between pretense and belief. In *Imagination, Philosophy, and the Arts*, M. Kieran and D. Lopes (eds.). London: Routledge.
- Goldman, A. 1989: Interpretation psychologized. *Mind and Language* 4, 161-185.
- Goldman, A. 1992a: Empathy, mind, and morals. *Proceedings and Addresses of the American Philosophical Association*, 66, 17-41.
- Goldman, A. 1992b: In defense of the simulation theory. *Mind and Language* 7, 104-119.
- Goldman, S., Graesser, A., van den Broek, P. (eds.) 1999: *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gordon, R. 1986: Folk psychology as simulation. *Mind and Language*, 1, 158-170.

- Gordon, R. and Barker, J. 1994: Autism and the 'theory of mind' debate. In *Philosophical Psychopathology: A Book of Readings*, G. Graham & G. L. Stephens (eds.). Cambridge, MA: MIT Press.
- Harris, P. 1991: The work of the imagination. In *Natural Theories of Mind*, A. Whiten (ed.). Oxford: Blackwell.
- Harris, P. 2000: *The Work of the Imagination*. Blackwell.
- Kunda, Z. 1999: *Social Cognition*. Cambridge, MA: MIT Press.
- Lang, P. 1984: Cognition and emotion: Concept and action. In *Emotions, Cognition and Behavior*, C. Izard, J. Kagan, R. Zajoncs (eds.). Cambridge: Cambridge University Press.
- Lang, P., Levin, D., Miller, G., and Kozak, M. 1983: Fear behavior, fear imagery, and the psychophysiology of emotion: The problem of affective response integration. *Journal of Abnormal Psychology*, 92, 276-306.
- Leslie, A. 1987: Pretense and representation: The origins of 'theory of mind'. *Psychological Review*, 94, 412-426.
- Leslie, A. 1994: Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211-238.
- Meskin, A. and Weinberg, J. 2003: Emotions, fiction, and cognitive architecture. *The British Journal of Aesthetics*.
- Moran, R. 1994: The expression of feeling in the imagination. *Philosophical Review*, 103, 75-106.
- Nichols, S. 2003: Imagination and the puzzles of iteration. *Analysis* 63.

- Nichols, S. 2004a: Imagining and believing: The promise of a single code. *Journal of Aesthetics and Art Criticism*.
- Nichols, S. 2004b: Review of G. Currie and I. Ravenscroft's *Recreative Minds: Imagination in Philosophy and Psychology*. *Mind*.
- Nichols, S. forthcoming: Imagination and immortality.
- Nichols, S. and Stich, S. 2000: A cognitive theory of pretense. *Cognition*, 74, 115-147.
- Nichols, S. and Stich, S. 2003: *Mindreading*. Oxford: Oxford University Press.
- Nichols, S., Stich, S., Leslie, A., and Klein, D. 1996: Varieties of off-line simulation. In *Theories of Theories of Mind*, P. Carruthers and P. Smith (eds.). Cambridge: Cambridge University Press.
- Noordman, L., Vonk, W., Kempff, H., 1992: Causal inferences during the reading of expository texts. *Journal of Memory & Language*, 31, 573-590.
- Sanitoso, R., Kunda, Z. and Fong, G. 1990: Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology*, 57, 229-241.
- van den Broek, P., Lorch, R., Linderholm, T., and Gustafson, M. 2001: The effects of readers' goals on inference generation and memory for texts. *Memory and Cognition*, 29, 1081-1087.
- Zwaan, R., Magliano, J., Graesser, A. 1995: Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 386-397.