



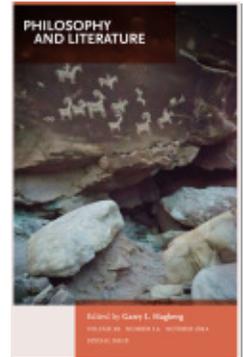
PROJECT MUSE®

Why the Novel Happened: A Cognitive Explanation

Tony Jackson

Philosophy and Literature, Volume 38, Number 1A, October 2014,
pp. A75-A93 (Article)

Published by The Johns Hopkins University Press
DOI: [10.1353/phl.2014.0025](https://doi.org/10.1353/phl.2014.0025)



➔ For additional information about this article

<http://muse.jhu.edu/journals/phl/summary/v038/38.1A.jackson.html>

III. The Emergence of the Novel, Evolutionary Psychology, and Literary Aesthetics

TONY JACKSON

WHY THE NOVEL HAPPENED: A COGNITIVE EXPLANATION

Abstract. Cognitively oriented literary studies, if they are to appeal to a broad array of literary scholars, will need to link cognition and culture. This essay brings together cognitive-psychological studies of the metarepresentational mind and of religious belief in order to offer an explanation of the nature and historical emergence of novelistic realism. It shows how novelistic realism, unlike other kinds of story, directly exercises what psychologist Alan Leslie calls the “decoupling mechanism” of the metarepresentational mind. And it argues that this kind of story takes on its specific power in the history of storytelling because of specific cultural change.

I

IN 1987, PSYCHOLOGIST ALAN Leslie published the essay “Pretense and Representation: The Origins of ‘Theory of Mind.’”¹ Even after more than twenty years, this remains a benchmark essay, having been cited over seven hundred times in the PsychINFO database as of summer 2011. “Theory of mind” is the cognitive-psychological term for the human ability to attribute mental states—intentions, desires, emotions—to

others. Our social being depends on this ability, which humans demonstrate from infancy, though, of course, it develops as the child grows. This ability depends on a certain kind of thought process that Leslie calls "metarepresentation." Leslie's essay offers an explanation of how metarepresentational thinking underlies a child's ability to understand other people, and so, to develop a theory of mind.

In psychological parlance, all conscious thoughts are representations. Primary representations, Leslie writes, "directly represent objects, states of affairs, and situations in the world" (pp. 416–17). Humans share this kind of representation with most other animals: a primary representation is our immediate awareness of the real world. But humans have a special ability to represent, not just aspects of the real world, but other representations as well. Said another way, we have the ability to think about what we are not immediately experiencing, and even to think about what has never been literally present to the perceiving mind. This latter case is what will most interest me here. A thought of something that we have not experienced is a representation, but cannot be a primary representation. Obviously, none of us have ever experienced someone else's thoughts or feelings, so our knowledge of such thoughts or feelings consists of metarepresentations. Theory of mind depends on our metarepresentational capacity.

Part of Leslie's goal in the essay is to explain how a metarepresentational capacity can possibly work, especially in the developing mind of a child. "The basic evolutionary and ecological point" of primary representation, he says, "must be to represent aspects of the world in an accurate, faithful, and literal way, in so far as this is possible for a given organism" (p. 414). In order to operate successfully in the world, any "perceiving, thinking organism ought, as far as possible, to get things right" (p. 412). But then the question arises about pretense and pretend play, which are not accurate, faithful, literal representations of the real world: "How is it possible for a child to think about a banana as if it were a telephone, a lump of plastic as if it were alive, or an empty dish as if it contained soap? If a representational system is developing, how can its semantic relations tolerate distortion in these more or less arbitrary ways? . . . Why does pretending not undermine [the child's] representational system and bring it crashing down?" (p. 412).

For Leslie, the issue is not just that we can produce metarepresentations, but that we can do so without wrecking the mind's primary cognitive function: accurate representation of the actual world in which we must survive. Leslie answers his question by arguing that a

“representational mechanism” must underlie the child’s ability to act as if the banana were a telephone while knowing that it is, in fact, a banana. This mechanism, also called a “decoupler,” ensures that metarepresentations are always systematically decoupled from primary representations (p. 417). In other words, even two-year-olds who pretend that the banana is a telephone consistently show awareness that the banana is always a banana. Leslie stresses the more or less mechanistic nature of all this (calling it a “processing subsystem” elsewhere²), because the evidence from very young children shows that humans arrive in the world predisposed to successfully perform metarepresentational thinking. It does not have to be learned. It is a part of what it means to be a human animal.

Though Leslie was explaining certain mental operations in terms of kinds of representations, it is not surprising that his work has gone on to become one of the founding texts in both the cognitive-psychological and the evolutionary-psychological study of the imagination.³ To metarepresent is, in everyday language, to imagine. This notion of imagination is hardly different from the dictionary definition (i.e., the ability to think of things not actually present to the senses). But if we can show that imagination is, in a key sense, natural—that it is a built-in, constitutive element of our cognitive architecture—then we open further avenues of explanation that might not otherwise appear to us. We can, for instance, think differently about beliefs in which the separation between primary representations (i.e., of the material world of experience) and metarepresentations (i.e., of the imaginary) is not carried out in the usual ways.

This brings us to the cognitive study of religious belief; or, more generally, belief in the supernatural. Religious belief has been taken up as a subject of cognitive research for only about the past twenty-five years. As recently as 2000, one of its primary investigators, Justin Barrett, could write an introduction to “the new cognitive science of religion” that was then “only about ten years old.”⁴ Since the 1990s, psychologists and cognitive anthropologists such as Barrett himself, Stewart Guthrie, Dan Sperber, Jesse Bering, Joseph Bulbulia, Robert McCauley and Harvey Whitehouse, Pascal Boyer, Scott Atran, and others have established a significant body of cognitively based explanations for various elements of religious belief and ritual.⁵ Though a cognitive-scientific approach to religious belief can possibly remain agnostic as to whether, say, supernatural agents actually exist, such an approach will necessarily try to explain belief in the supernatural as a function of naturally occurring processes. No one can dispute that such beliefs are extremely common

and important in all places and times, as magic, superstition, and more formalized forms of religion.

For these and other reasons, a cognitive approach assumes that such belief is not a function of culture alone (though any specific instance of religious belief will necessarily have its specific cultural inflections). Rather, on some important level, religious kinds of thought and behavior must be a function of the evolved human mind/body. Cognitive investigations, then, operate according to what “may be called the ‘naturalness-of-religion thesis.’”⁶

From early on, supernatural concepts have been explained in relation to another important set of established ideas, one having to do with our innate, intuitive knowledge of the world in which we have evolved. At this point, there are three categories or domains of intuitive knowledge that are widely accepted by cognitive psychology and anthropology: “intuitive biology, intuitive physics, and intuitive psychology.”⁷ Because this research is still young, the proper labels for these domains are not quite settled. So we sometimes read of “folk” biology, physics or mechanics, and psychology (another term for theory of mind); or, in the study of child development, “naïve” biology, physics, and psychology. (I shall stick with “intuitive” here.) But the definitions are consistent. As Atran explains:

Recent developmental, cognitive and crosscultural experiments strongly indicate that all (non-brain-damaged) humans have distinct core faculties of mind, with privileged access to these distinct but overlapping domains of nature: folk mechanics (object boundaries and movements), folk biology (biological species configurations and relationships), and folk psychology (interactive agents and goal-directed behavior).⁸

These core faculties are already clearly present in rudimentary form even in infants. Of course, such knowledge may not in every case be valid. For instance, our intuitive physics understands the sun as revolving around the Earth. But still, most of the time, our domain-specific intuitive knowledge serves us very well, giving us dependably valid information about the world and other people. If this were not the case, it is hard to see how we could have survived as a species.

But religious beliefs are very commonly counterintuitive, which here means not “strange, inexplicable, funny, exceptional or extraordinary” but rather in violation of our innate, intuitive understandings.⁹ In fact, our three most established kinds of intuitive knowledge provide a

template for analysis of the nature of any specific supernatural concept. We can explain any given supernatural concept or entity by showing the specific ways in which it runs “counter” to our domain-specific, intuitive understandings. Boyer gives an example from an Andean people, the Aymara, who think of a local mountain as a living creature, with body, head, lungs, the ability to eat, bleed, and so on. An entire corpus of cultural thought and action is built around “this transfer of biological properties, associated with animals, plants, and persons, to what is otherwise identified as an inert natural object.”¹⁰ In general, human beings simply know, intuitively, that mountains do not fall within the realm of living things, so this is a case of counterintuitive biological belief. This Aymara example is simple and straightforward, but it is paradigmatic for other, more complicated counterintuitive beliefs, in which a given entity might violate, for instance, both intuitive biological and intuitive psychological understandings of the world.

Now, counterintuitive concepts are hardly limited to religious kinds of thinking. As I will discuss at more length below, they show up all the time in myth, fantasy, and science fiction. But, of course, the difference is that in the case of religion we have counterintuitive *beliefs*, not just counterintuitive *concepts*. To quote Atran:

One significant distinction between fantasy and religion is knowledge of its source. People know or assume that public fictions (novels, movies, cartoons, etc.) were created by specific people who had particular intentions for doing so. Religious believers, however, assume that the utterances or texts connected with religious doctrines are authorless, timeless, and true.¹¹

We imagine counterintuitive concepts all the time, but, except for religious kinds of belief, our decoupling mechanism prevents us from (mis)taking such concepts as “veridical beliefs”: beliefs grounded in both established beliefs and material evidence based on accurate appraisal of the world in which we have evolved.¹²

It might seem to a skeptic that once *any* counterintuitive concept is taken as a veridical belief, then anything goes—*anything* may be believed in any way, no matter how impossible. And the sheer variety of religions and superstitions around the world appears to support this idea. But counterintuitive concepts in general, and religious beliefs in particular, are not simply unconstrained flights of fantasy. They operate in predictable ways that are dependent on human memory and attention. Research has shown that any stories with “violations of intuitive

expectations” are much more likely to be recalled than stories without such violations.¹³ This is now a commonly accepted explanation of why religious concepts are so readily taken up by so many people. Just because their content is, strictly speaking, impossible, counterintuitive concepts are more likely to grab our attention, and are easier to remember than other kinds of concepts.

At the same time, however, “massively counterintuitive” ideas won’t work.¹⁴ No matter how “superstitious” a person or culture might be, supernatural concepts that are overburdened with too many counterintuitive properties will be seen as nonsense. Counterintuitive concepts must be inconsistent with our intuitive understandings in order to be memorable; but the inconsistency must be minimal, because too much impossibility becomes harder to recall. Indeed, it turns out that concepts involving only one intuitive violation are the most memorable of all.¹⁵ And the impossible elements of a given belief are most commonly cast against a more or less everyday, intuitively dependable material-world background. In the example mentioned by Boyer, the Aymara “do not have a fantastic ontology in which mountains *in general* are live organisms. . . . The supernatural concept specifies that *this* mountain has *some* physiological features.”¹⁶ All other mountains are treated like mountains in the usual ways. In fact, the focal counterintuitive concept requires the everyday, taken-for-granted, intuitive background in order to take on its special meanings.

With these most basic elements of religious belief laid out, I now return to the broader issue of the metarepresentational mind. Clearly, any religious belief depends on metarepresentation. Our “metarepresentational capacity . . . is common to everyday cognition and communication and not particular to religion. Without it, though, representations of what gods are or do, and the entertaining of other religious quasi propositions, would not be possible.”¹⁷ But really, from the cognitive- and evolutionary-psychological perspectives, the distinctive quality about religious beliefs is that they are a case—in fact, the great case—of the metarepresentational mind systematically failing to decouple the imaginary from the real.

And yet, as mentioned above, this kind of belief must be a function of the way the metarepresentational mind operates. Religious beliefs have been universal in human culture for an extremely long time, and “probably appeared with the modern mind.”¹⁸ McCauley argues that, unlike the task of learning a scientific discipline, learning religion is

easy; it is akin to learning a natural language, because, like “natural language, religion exploits cognitive dispositions, which seem to arise early in human development.”¹⁹ Research suggests that children, from a very young age, understand the world as determined by a godlike teleology, and that adults, even scientists, often do the same, apart from any specific religious belief.²⁰ Guthrie, in *Faces in the Clouds: A New Theory of Religion*, explains at length how we generate religious beliefs automatically as a function of our built-in tendency to anthropomorphize the world. Bering, a psychologist, writes that “a general belief in the continuity of mental states in dead agents seems not something that children acquire as a product of their social-religious upbringing. . . . Instead, a natural disposition toward afterlife beliefs is more likely the default cognitive stance and interacts with various learning channels” encountered in cultural existence.²¹

There is much evidence, then, to show that religious (counterintuitive) beliefs are in general a function of the metarepresentational mind. Such beliefs must, in their unique way, be “natural.” We are describing a cognitive function, so in this case, “natural” does not mean rigidly constrained, as with a purely physiological function. Rather, it means (in Bering’s view) a disposition toward a certain kind of behavior or thinking that will take on its specific, manifest expression as a function of appearing in a specific cultural context (Bering’s “learning channels”). If this cognitive function is “natural” in this way, then we may have reasonable expectations of how it might fare in differing cultural situations. In other words, though cultural contexts will vary according to place and time, the underlying cognitive function will not. This is a founding concept for all cognitive-psychological, as well as all evolutionary-psychological, studies of the arts.

A quick parallel with evolutionary psychology may be helpful here. In *The Art Instinct*, Denis Dutton made this kind of evolutionary psychological case for the appeal of art in general. In order to give a Darwinian basis to the human interest in imaginative creation, Dutton combined an array of indisputably cross-cultural instances of music, visual art, and storytelling; a judicious understanding of evolutionary theory; and constant attention to what can legitimately be considered universal, as opposed to culturally specific. From this basis, he could then claim for narrative fiction (as an example) that, if “there was adaptive survival value in ancient, Stone Age storytelling, it ought to extend to our own time and explain somehow the pleasure we get from any fictions.”²²

The same is true for the kinds of cognitively based universals I am considering here. These universals give us an empirically established basis for understanding specific cultural artifacts, practices, and beliefs. For instance, if cultural changes make it difficult to maintain counterintuitive concepts as veridical beliefs, then we can expect in all likelihood to find the cognitive function still operating—but in other ways.

II

Being a student of narrative, I will consider the “other ways” as they occur in the realm of storytelling. This makes sense, first, because of the close ties between religion and storytelling. Many, many religious beliefs are ultimately anchored in story and myth: strictly speaking, there is no other anchor to be had for a counterintuitive concept. But, second, I turn to storytelling because of what I mentioned above in passing: fantasy, by which I mean fictions involving counterintuitive concepts, as opposed to counterintuitive beliefs. Defined this way, fantasy encompasses all those stories, new and old, that involve ghosts, dragons, metamorphoses, one-eyed monsters, talking mirrors or trees or animals or plants, witches, sorcerers, frog-princes, mutant humans, superheroes, enchanted forests, demonic cars, any kind of magic, et cetera, et cetera. A complete list would be very long indeed, and would have examples from all languages and times. The pervasive presence and extremely wide appeal of such stories can lead us to at least wonder if they may reasonably be considered a kind of “default” human story.

In any event, what matters here is that fantasy stories feature counterintuitive content that is experienced as fiction, which means our decoupling mechanism is working in the standard way: to ensure that we do not confuse the veridical with the imaginary impossible. *Within* the story-world of a fantasy, the counterintuitive content typically operates in the way that the content of religious beliefs operates in everyday human life: as uniquely special cases, yet also as part of the material world in which we operate. But readers (or audiences) do not take the stories as nonfiction.

And yet, having made this last point, I have overgeneralized. For, in fact, the distinction between religious belief and fantasy tends to be a modern, literate one. Primary oral cultures, which have been uninfluenced by literacy, do not separate myth, religion, history, and folklore from one another in the way that scientific, hyperliterate cultures do. Cultural changes associated with modernity have caused this distinction

to happen. As is well known, in the modern era both common “superstitious” belief and institutionalized religious belief have faced unprecedented, continual, strong opposition, both directly and indirectly, from scientific and related, hyperskeptical kinds of thinking. As Jean-Pierre Changeux has argued, such beliefs “have to a considerable extent been transformed or replaced in the course of recent history by . . . scientific ideas.”²³ So cultural changes have forced distinctions between fantasy, myth, history, and religion that we moderns now tend to take for granted.

But why has the inexorable increase in the power of scientific and other rationalist ideas since the time of Francis Bacon not simply vanquished religious beliefs altogether? The explanation is that the default cognitive function discussed above continues to operate. This fact can help explain why no amount of scientific knowledge, political coercion, or socioeconomic affluence has been able to eliminate even the most conspicuously counterintuitive religious kinds of belief. Counterintuitive concepts still have a fundamental appeal, evident not only in the continuing success of religion, but also in the continuing, very widespread popularity of fantasy as a distinctive kind of story. If the metarepresentational mind finds its counterintuitive *beliefs* relegated by cultural change to the realm of nonsense or childish superstition, it can still preserve counterintuitive *concepts* in the form of the fantasy story.

Broadly considered, science fiction is the most conspicuous example of a modern preservation and transformation of fantasy. We could reasonably expect narratives called “science fiction” to be distinctly *not* fantasy, *not* based on magic or supernatural phenomena, but rather somehow based on science. However, what we very often (though by no means always) get are counterintuitive concepts that are made acceptable (i.e., not magical) because they are explained in terms of some version of scientific possibility.

The spectacular emergence of the superhero in the last half century is exemplary. For instance, Marvel Comics’ *Fantastic Four* features four characters who are exposed to cosmic rays and, as a result, develop superpowers. Exotic “rays,” because they are linked with scientific discoveries, take the place of various kinds of magical metamorphoses in myth and legend. As with religious concepts, the superpowers of the four involve distinctly limited impossibilities: one can stretch his body into extreme shapes, one can become invisible, one can generate flames, and one has superhuman strength and a body made of stone.

Apart from their enemies’ equally limited impossibilities, the *Fantastic Four* story-world is otherwise our everyday world. But, as is common

with science fiction, that world is staged one way or another in a future time. Situating science-fiction stories in the future makes counterintuitive impossibilities all the more acceptable. Staging stories somewhere away from Earth often enhances the same effect. I do not suggest that all science-fiction stories operate in these ways, only that many of the classic stories do. This generic placing of the story in the future and far away has its exact parallels in the most ancient forms of fantasy. Legends, folk tales, creation myths, and the like regularly make their counterintuitive content more acceptable by locating the story in the distant past and/or in a distant land. When director George Lucas begins the *Star Wars* science-fiction saga with “A long time ago in a galaxy far, far away,” he is making sure to capture both the most ancient and the most contemporary fascination with fantasy.

Thus, we can see that, with the emergence of scientific kinds of thinking—thinking that most seriously undermines any counterintuitive content—the metarepresentational mind still produces, and is fascinated by, counterintuitive concepts, though adjustments must be made in order to make those concepts as conformable as possible to current cultural contexts. As Lisa Zunshine puts it, counterintuitive stories appeal to us, now as in the past, because they “tease in particularly felicitous ways” specific kinds of cognitive functions.²⁴

But, if counterintuitive concepts have such a fundamental appeal to the metarepresentational mind, what are we to think of the kind of storytelling that famously establishes its generic identity by, among other things, forthrightly rejecting counterintuitive concepts? Any student of literature will immediately know that I am referring to the novel. It is the novel as a genre that most concerns me here. Though the novel may reject counterintuitive concepts, its nature may still be explained through the cognitive concepts laid out above. To demonstrate this, I will first offer a general, cognitively oriented explanation for the appeal of the novel, and then turn to some specific examples.

The emergence of the novel in the seventeenth and eighteenth centuries marks a significant event in the history of storytelling, in part because the novel, as a genre, has a distinctive commitment to what we now call realistic storytelling—no counterintuitive agents or events allowed. Of course, there are examples of long prose fiction before the seventeenth century; and, of course, there are examples of realistic stories before the seventeenth century. But then, it is rare to discover a historically new entity that simply has no precedents. The extraordinary success of the novel as a genre distinguishes it from the cases of long prose fiction or

realistic stories that preceded it. Furthermore, there are examples of long prose fiction after the seventeenth century that include all manner of counterintuitive content. This is to be expected. In much the same way that “no development in science will ever seriously threaten the persistence of religion,”²⁵ no development in storytelling will ever seriously threaten the persistence of stories with counterintuitive content.²⁶

Despite this, if we are to mark off the novel as a significant entry in the history of storytelling, we do so in large part because of its historically distinctive attempt “to make literature appear to be describing directly . . . reality itself,” to “convey the impression of an empirically shareable experience.”²⁷ Said another way, even though all stories that operate as fiction are a kind of pretend play, novelistic realism stands out historically because it wants to be “maximally reproductive of [the] world it is modeling for play purposes.”²⁸ Though that empirical world may include people who have counterintuitive beliefs, it cannot include counterintuitive impossibilities themselves.

Novelistic realism has a paradoxical appeal. We know that we are reading a consciously created fiction, and yet the story strikes us much in the way of a veridical report of the everyday life of everyday people. On a most fundamental level, our enjoyment stems from the experience of a fiction coming as close as possible to nonfiction. To return to our cognitive terms: this kind of story is a consciously created metarepresentation that strives to convey the uncreated experience of primary representation, and our enjoyment arises from just this fact.

Said another way: a defining characteristic of the novel as a historically important kind of story is that it makes the *cognitive* distinction (as opposed to a metaphysical, epistemological, or ontological distinction) between imagined fiction and a report of the real a featured element of the reader’s experience. Novelistic realism, then, is fiction directed toward our “decoupler,” the mechanism by which we ordinarily distinguish metarepresentation from primary representation. This becomes clear if we consider in broad terms how novelistic realism differs generically from fantasy. Since it is impossible to have an actual counterintuitive experience, no fantasy story can have the feel of a primary representation in the way that novelistic realism strives for. With counterintuitive concepts as they occur in fantasy and counterintuitive beliefs as they occur with regard to religion, the decoupling mechanism is either working (fantasy) or not working (religion), but is not itself at issue. Neither fantasy nor religious belief wants to foreground, so to speak, the mechanism that enables their successful operation. In contrast, as

a kind of story, novelistic realism directly exercises, we might say, our decoupling mechanism.

This claim is also supported by a consideration of literary or creative nonfiction, which operates as an intimately related mirror opposite of novelistic realism. The content of literary nonfiction appeals because it is based on specific, uncreated primary representations and yet strives to read as much like created fiction as possible. The story must involve documentable content taken from the real world as opposed to being simply invented, and commonly involves extensive research in order to secure the referential fidelity to the real world. At the same time, it must have a literary style such that it distinctly does not read as objective, journalistic reportage. If just the nonfictional content were the main appeal, then a biography or a history or a documentary report would satisfy.

But that is not the case. Literary nonfiction strives to read like a novel, complete with the kinds of descriptive detail, conversation, reported thoughts, and dramatic excitement that are seldom possible with a documentary fidelity to the facts. And yet, if only the “literary” or “creative” element were what a reader was looking for, then a novel would satisfy. As with novelistic realism, literary nonfiction makes the cognitive distinction between imagined fiction and documentary report, between metarepresentation and primary representation—a featured element of the reader’s experience.

The intimate relationship between novelistic realism and literary nonfiction has been present since the novel’s beginnings. For instance, in the prologue to *The Adventures of Don Quixote*, Cervantes speaks of his book as his own imaginary creation, “the child of my brain.”²⁹ Nonetheless, in the book itself he consistently presents himself as merely the “second author” of an already written documentary history of “our most famous Spaniard.” In the middle of the adventure of the windmills, the story suddenly stops—because, we are told, the original “author of this history left the battle in suspense at this critical point, with the excuse that he could find no more records of Don Quixote’s exploits than those related here.”³⁰ Luckily, the rest of the written record, authored by one Cid Hamete Benengeli, turns up in Quixote’s village, so the story can continue. Throughout the novel, Cervantes is ironic about the actuality of the “second author” and, indeed, about the writing of this kind of “history.” But still, he plainly felt the need to present the tale as literary nonfiction. (I will examine this paradigmatic text in more detail below.)

Other examples show this same authorial concern. Daniel Defoe, on the title page of *Robinson Crusoe*, lies outright and tells us this is “The life” of its author as “Written by himself.” Then, in the preface, the “Editor” claims in a defensive tone that the tale is “a just History of Fact; neither is there any Appearance of Fiction in it.”³¹ But it turns out that this “history of fact” is only interesting because it involves “Strange Surprising Adventures.” Defoe sensed, correctly, that the strange, surprising adventures, if presented only as fiction, would not grab an audience as would a factual life that reads like an adventure story. Aphra Behn, in the dedication of *Oroonoko*, claimed that it was “a true story” with nothing “Romantick” in it, though she also warned (and so, enticed) her reader that tales from the tropical “other world” would naturally involve “unconceivable wonders” and seem “New and Strange” to Europeans.³² The writer of the epistolary novel *Letters of a Portuguese Nun*, published anonymously in 1669, went to the extreme of never revealing that the nun’s passionate love letters were fictions. And, more generally, the epistolary novel, which was so essential in the emergence of the novel as a genre, typically presented itself as nonfiction. Samuel Richardson, for instance, claimed to be merely the editor of the batch of “real” correspondence collected in *Pamela*, and continued to make such claims even when the public at large knew he was the sole author.

So in order to gain entrance into the house of storytelling, the novel had to first establish its generic place by pretending to be literary nonfiction. We may explain the historical emergence of novelistic realism in much the same way we have explained the continued existence of fantasy in modern times, though the case of the novel is more complex. As Blakey Vermeule has explained in cognitive terms, “Cultural forms adapt to fulfill deep psychological imperatives.”³³ The novel as a genre is such an adaptation. Novelistic realism makes its mark in the history of storytelling along with “the development of the belief that the new scientific methodologies were the only reliable measure of an external reality.”³⁴ Counterintuitive concepts in general begin to be downgraded by scientific rationalism (and other, related forces) to the level of childish wish fulfillment, mindless superstition, or naïve idealism. If cultural changes cause counterintuitive concepts to be ever more difficult to sustain, then we may expect the cognitive system that has underwritten those concepts to continue to operate, but that other cultural forms will become its input.

We can draw a very rough comparison here with bodily exercise. We must assume that the human body evolved as a function of the kinds

of physical activities necessary to ensure our survival. Such activities would have occurred simply as a function of living everyday life. Thus, hunter-gatherer and agricultural societies have no need for what we call “exercise” because living everyday life *is* exercise. But once cultural changes remove the need for such everyday activities, we come up with “exercise.” As religious belief and its pretend form in fantasy become less and less efficacious as a result of cultural change, novelistic realism (and literary nonfiction) emerges as a kind of direct exercise of our decoupling mechanism. If we cannot have the belief itself, we can still exercise the mechanism that enables that belief. The difference, of course, is that in this case we are dealing with unconscious rather than conscious productions of alternative kinds of stimulation.

III

It makes sense, then, that one of the defining markers in the history of storytelling—*Don Quixote*—would take Quixote’s peculiar madness as its subject. In cognitive terms, novelistic realism is always, in a sense, “about” the demotion of counterintuitive content from its ancient, unquestioned preeminence in the house of story. *Quixote* inaugurates novelistic realism by directly performing this demotion. With our cognitive understanding in mind, let us look a bit more closely at this great novel.

Commonly considered one of the first novels, Cervantes’s fine tale is widely known and extraordinarily influential in the history of literature in general and the history of the novel in particular.³⁵ A Spanish aristocrat loses his mind by obsessively reading chivalric romances: the fantastic medieval tales of knights, damsels, demons, sorcerers, and magical transformations. He comes to (mis)believe that he himself is a knight, and that the everyday world is the scene of all manner of fantastic occurrences. Taken as a whole, the novel has often been read as a parody of—and, in the end, even a polemical rejection of—the romance tales. In this way, Cervantes began to establish the novel as a realistic genre by lampooning stories with counterintuitive content. Though the tale may not seem realistic in our post-nineteenth-century sense, it remains the case that “no one before [Cervantes] had infused the element of genuine everyday reality” into this kind of adventure tale.³⁶ Even scholars who do not want to count *Quixote* as a distinct case of novelistic realism nonetheless admit that it “reveals the dawn of realism on the horizon.”³⁷ Also, the story must be fundamentally realistic simply

in order to portray Quixote's madness as it does. And this portrayal of madness is rich in meaning.

I have mentioned that religious beliefs and fantastic stories have specific, limited counterintuitive elements cast against a more or less everyday background. As is common with religion and fantasy, *Quixote* features just one major element that has to do with the counterintuitive: its main character systematically fails to make the usual distinction between primary representations and metarepresentations. He regularly experiences perceptual actualities as counterintuitive impossibilities. But the impossibilities are clearly limited in kind: all are dependent on the traditions of chivalric romance. They are, however, limited in other ways as well. Quixote does not see the world as entirely a chivalric romance (as the Aymara do not see all mountains as living things); only certain primary representations are mistaken as metarepresentations. Otherwise, he moves and acts in the world as does everyone else. And the world around him operates in the usual ways: no counterintuitive events occur in the novel.

For instance, Quixote (mis)believes that an inn is a castle and the innkeeper is its lord. He transforms the peasant girl, Aldonza Lorenzo, into the fantastic Dulcinea, an impossibly perfect royal beauty. He sees a barber's brass basin as the legendary, magical Mambrino's helmet. The list is substantial. But perhaps the single most famous example occurs when Quixote (mis)believes that windmills are giants, and attacks them forthwith. This image has long been the defining action of the larger cultural icon that the figure of Don Quixote has become. For, even though the book may be the "most widely read literary masterpiece in world literature," its fame is not limited to its readers.³⁸ People who have never read the novel will very often know of the crazy old man who roams the countryside tilting at windmills. Adaptations of the story make sure to include it; most famously, in the 1965 Broadway musical *Man of La Mancha* (made into a feature film in 1972).

Scholars puzzle over why this image should be so widely known.³⁹ But one clear reason is that, by attributing biological properties to a non-biological entity, it operates cognitively just in the way of other limited, counterintuitive concepts. The nature of such concepts is that they are especially attention grabbing and easily remembered, even when they occur in a parody. But in this case most memorably of all, we have, not a counterintuitive concept or belief, but a man who fully believes in the material existence of his counterintuitive concepts.

With our cognitive perspective in mind, then, we can see that this originary example of novelistic realism focuses directly on a character whose decoupling mechanism systematically fails to separate counter-intuitive metarepresentations from primary representations, which is part of what makes Quixote so interesting and memorable. Consciously created fantasy stories (the medieval romances) operate for Quixote in the same way that religious stories typically operate for true believers. This parallel is reinforced by the way he remains committed to his (mis) beliefs. Because of the way that the mind holds religious beliefs, they cannot be disconfirmed; in fact, commonly, the more the world seems to disconfirm them, the stronger the true believer's beliefs grow.⁴⁰ In the end, Quixote does give up his counterintuitive beliefs, but along the way they are regularly strengthened by disproving evidence and opinion—even though he is otherwise a quite rational person.

But, of course, nowhere in the novel is any connection drawn between Quixote's madness and religious belief. Cultural context may have enabled the production and enjoyment of the image of a man who fails to decouple metarepresentation from the real, but that same context disallows extending that image to religious belief itself. (This is not to suggest that religious belief is somehow a form of madness: it is not, as the cognitive psychology of religious belief plainly shows.) Such extensions could only come much later in history.

All examples of novelistic realism will directly exercise the mechanism by which we distinguish metarepresentation from primary representation, while only some will be about characters whose ability to decouple imaginary from real fails as it does for Quixote. As has long been noticed, just such characters have regularly marked changes in the history of novelistic realism. *The Female Quixote* (1752), *Waverley* (1814), *Northanger Abbey* (1818), and *Madame Bovary* (1856), among others, are novels about a main character who fails to decouple his or her reading of fiction from the material actualities of lived life. Along with this, we may read the history of novelistic realism as a history of innovations that work to cause the story to be taken as qualitatively more like a report of everyday lived experience. Each new marker in the history of the novel signals an advance "in fictional technique by freeing itself from the conception of the real [that is] currently dominant" and moving on toward "real" realism.⁴¹ We will evidently need to constantly develop new means of exercising this mental mechanism. Given all this, and with *Don Quixote* as our paradigm, it seems at least likely that my claims about

the metarepresentational mind and its stories lay out the possibility of a cognitively oriented reading of the history of the novel as a genre.

UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE

1. Alan Leslie, "Pretense and Representation: The Origins of 'Theory of Mind,'" *Psychological Review* 94 (1987): 312–426. Subsequent references are inserted in the text.
2. Ori Friedman, Karen Nearya, Corinna Bunnstein, and Alan Leslie, "Is Young Children's Recognition of Pretense Metarepresentational or Merely Behavioral? Evidence from 2- and 3-Year-Olds' Understanding of Pretend Sounds and Speech," *Cognition* 115 (2010): 314.
3. See, for instance, *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*, ed. Shaun Nichols (Oxford: Oxford University Press, 2006); Gregorie Currie and Ian Ravenscroft, *Recreative Minds: Imagination in Philosophy and Psychology* (Oxford: Oxford University Press, 2002); Leda Cosmides and John Tooby, "Consider the Source: The Evolution of Adaptations for Decoupling and Metarepresentation," in *Metarepresentations: A Multidisciplinary Perspective*, ed. Dan Sperber (Oxford: Oxford University Press, 2000), pp. 53–115; and Denis Dutton, *The Art Instinct: Beauty, Pleasure, and Human Evolution* (New York: Bloomsbury Press, 2009).
4. Justin L. Barrett, "Exploring the Natural Foundations of Religion," *Trends in Cognitive Sciences* 4 (2000): 29.
5. Stewart Guthrie, *Faces in the Clouds: A New Theory of Religion* (Oxford: Oxford University Press, 1993); Stewart Guthrie, "Why Gods? A Cognitive Theory," in *Religion in Mind: Cognitive Perspectives on Religious Belief, Ritual, and Experience*, ed. Jensen Andresen (Cambridge: Cambridge University Press, 2001), pp. 94–112; Dan Sperber, *Explaining Culture: A Naturalistic Approach* (Oxford: Blackwell, 1996); Jesse Bering, "The Folk Psychology of Souls," *Behavioral and Brain Sciences* 29 (2006): 453–98; Joseph Bulbulia, "The Cognitive and Evolutionary Psychology of Religion," *Biology and Philosophy* 19 (2004): 655–86; Robert N. McCauley and Harvey Whitehouse, "Introduction: New Frontiers in the Cognitive Science of Religion," *Journal of Cognition and Culture* 5 (2005): 1–13; Pascal Boyer, *The Naturalness of Religious Ideas: A Cognitive Theory of Religion* (Berkeley: University of California Press, 1994) and *Religion Explained: The Evolutionary Origins of Religious Thought* (New York: Basic Books, 2001); Scott Atran, "Causal Constraints on Categories," in *Causal Cognition: A Multidisciplinary Debate*, ed. Dan Sperber, David Premack, and Ann James Premack (Oxford: Oxford University Press, 1995), pp. 205–33, and *In Gods We Trust: The Evolutionary Landscape of Religion* (Oxford and New York: Oxford University Press, 2002). For useful book-length surveys of the cognitive psychology of religious belief, see Todd Tremlin, *Minds and Gods: The Cognitive Foundations of Religion* (Oxford: Oxford University Press, 2002); and Ilkka Pyysiäinen, *Supernatural Agents: Why We Believe in Souls, Gods, and Buddhas* (Oxford: Oxford University Press, 2009). For a short but precise overview, see Pascal Boyer's 2008 essay "Religion: Bound to Believe?" *Nature* 455 (2008): 1038–39.

6. Barrett, "Exploring the Natural Foundations of Religion," p. 29.
7. Tremlin, *Minds and Gods*, p. 66.
8. Scott Atran, "Strong versus Weak Adaptationism in Cognition and Language," in *The Innate Mind: Structure and Contents*, ed. Peter Carruthers, Stephen Laurence, and Stephen Stich (Oxford: Oxford University Press, 2005), p. 143.
9. Boyer, *Religion Explained*, p. 65. See also Tremlin, *Minds and Gods*, pp. 66–68; Pyysiäinen, *Supernatural Agents*, pp. 22–30; and Atran, *In Gods We Trust*, pp. 83–113.
10. Boyer, *Religion Explained*, p. 66.
11. Atran, *In Gods We Trust*, p. 267.
12. Ryan McKay and Daniel Dennett, "The Evolution of Misbelief," *Behavioral and Brain Sciences* 32 (2009): 494.
13. Boyer, *Religion Explained*, p. 80. See also Atran, *In Gods We Trust*, pp. 100–107; and Ara Norenzayana, Scott Atran, Jason Faulkner, and Mark Schallera, "Memory and Mystery: The Cultural Selection of Minimally Counterintuitive Narratives," *Cognitive Science* 30 (2006): 531–53.
14. Pyysiäinen, *Supernatural Agents*, p. 27.
15. See Boyer, *Religion Explained*, p. 86; Pyysiäinen, *Supernatural Agents*, pp. 26–28; Tremlin, *Minds and Gods*, p. 90; Atran, *In Gods We Trust*, pp. 106–13; and Barrett, "Exploring the Natural Foundations of Religion," p. 30.
16. Boyer, *Religion Explained*, p. 66.
17. Atran, *In Gods We Trust*, p. 108.
18. Boyer, *Religion Explained*, p. 322.
19. Robert McCauley, "The Naturalness of Religion and the Unnaturalness of Science," in *Explanation and Cognition*, ed. Frank C. Keil and Robert A. Wilson (Cambridge: The MIT Press, 2000), p. 80.
20. See Susan Gelman, *The Essential Child: Origins of Essentialism in Everyday Thought* (Oxford: Oxford University Press, 2003); Barrett, "Exploring the Natural Foundations of Religion"; Bulbulia, "The Cognitive and Evolutionary Psychology of Religion"; Deborah Kelemen, "The Scope of Teleological Thinking in Preschool Children," *Cognition* 70 (1999): 241–72, and "British and American Children's Preferences for Teleo-Functional Explanations of the Natural World," *Cognition* 88 (2003): 201–21; Deborah Kelemen and Evelyn Rosset, "The Human Function Compunction: Teleological Explanation in Adults," *Cognition* 111 (2009): 138–43; and Atran, "Causal Constraints on Categories."
21. Bering, "The Folk Psychology of Souls," p. 454.
22. Dutton, *The Art Instinct*, p. 126.
23. Jean-Pierre Changeux, *The Physiology of Truth: Neuroscience and Human Knowledge*, trans. M. B. DeBevoise (Cambridge: Harvard University Press, 2004), p. 220.
24. Lisa Zunshine, *Why We Read Fiction: Theory of Mind and the Novel* (Columbus: Ohio State University Press, 2006), p. 53.

25. McCauley, "The Naturalness of Religion and the Unnaturalness of Science," p. 61.
26. Tony E. Jackson, *The Technology of the Novel: Writing and Narrative in British Fiction* (Baltimore: Johns Hopkins University Press, 2009), p. 18.
27. George Levine, *The Realistic Imagination: English Fiction from Frankenstein to Lady Chatterley* (Chicago: University of Chicago Press, 1981), pp. 8, 18.
28. Peter Brooks, *Realist Vision* (New Haven: Yale University Press, 2005), p. 2.
29. Miguel de Cervantes Saavedra, *The Adventures of Don Quixote*, trans. J. M. Cohen (New York: Penguin Books, 1950), p. 25.
30. Cervantes Saavedra, *Don Quixote*, pp. 74–75.
31. Daniel Defoe, *Robinson Crusoe*, 2nd ed., ed. Michael Shinagel (New York: W. W. Norton, 1994), p. 3.
32. Aphra Behn, *Oroonoko; or, The Royal Slave*, ed. Catherine Gallagher (Boston: Bedford/St. Martin's Press, 2000), p. 37.
33. Blakey Vermeule, *Why Do We Care about Literary Characters?* (Baltimore: Johns Hopkins University Press, 2009), p. 10.
34. Edward Dudley, *The Endless Text: "Don Quixote" and the Hermeneutics of Romance* (Albany: State University of New York Press, 1997), p. 3.
35. See, for instance, Harry Levin, *The Gates of Horn: A Study of Five French Realists* (Oxford: Oxford University Press, 1963); Georg Lukács, *The Theory of the Novel: A Historico-Philosophical Essay on the Forms of Great Epic Literature*, trans. Anna Bostock (Cambridge: The MIT Press, 1971); and Manuel Durán and Fay R. Rogg, *Fighting Windmills: Encounters with "Don Quixote"* (New Haven: Yale University Press, 2006).
36. Eric Auerbach, *Mimesis: The Representation of Reality in Western Literature*, trans. Willard Trask (Princeton: Princeton University Press, 1968), p. 355.
37. Félix Martínez-Bonati, "Don Quixote" and the Poetics of the Novel, trans. Dian Fox (Ithaca: Cornell University Press, 1991), p. 193.
38. Durán and Rogg, *Fighting Windmills*, p. 1.
39. See Dudley, *The Endless Text*, p. 170; Robert Bayliss, "What *Don Quixote* Means (Today)," *Comparative Literature Studies* 43 (2006): 382–97; and E. C. Riley, "Don Quixote: From Text to Icon," *Cervantes* 8 (1988): 103–16.
40. Atran, *In Gods We Trust*, pp. 92–93.
41. Levine, *The Realistic Imagination*, p. 72.