HOW AUTHORS' MINDS MAKE STORIES

This book explores how the creations of great authors result from the same operations as our everyday counterfactual and hypothetical imaginations, which cognitive scientists refer to as "simulations." Drawing on detailed literary analyses as well as recent research in neuroscience and related fields, Patrick Colm Hogan develops a rigorous theory of the principles governing simulation that goes beyond existing frameworks. He examines the functions and mechanisms of narrative imagination, especially the role of theory of mind, and relates this analysis to narrative universals. In the course of this theoretical discussion, Hogan explores works by Austen, Faulkner, Shakespeare, Racine, Brecht, Kafka, and Calvino. He pays particular attention to the principles and parameters defining an author's narrative idiolect, examining the cognitive and emotional continuities that span an individual author's body of work.

Patrick Colm Hogan is a professor in the Department of English and the programs in Cognitive Science, Comparative Literature and Cultural Studies, and India Studies at the University of Connecticut. He is the author of fifteen books, including *The Mind and Its Stories* and *What Literature Teaches Us about Emotion*, and the editor of *The Cambridge Encyclopedia of the Language Sciences*.

How Authors' Minds Make Stories

PATRICK COLM HOGAN

University of Connecticut

CAMBRIDGE
UNIVERSITY PRESS

For Keith and David

Contents

Acknowledgments

Introduction

From the Universal to the Particular

Readers of Rabindranath Tagore's *My Reminiscences* may recall the importance of the blue manuscript book the author received as a young boy. With the encouragement of his older brother, he wrote and wrote in the book. That practice in written self-expression was the seed from which Tagore's later poetry grew.

In his short story "Exercise-Book," Tagore seems to take up some of this personal history. Uma, the main character, also writes feverishly in a notebook. Indeed, the notebook serves as the major outlet for this character's feelings and thoughts, just as it did for young Rabindranath. But there is a difference. This hero is a girl. Rather than being celebrated by her older brother, she is married off at the age of nine (in keeping with Bengali customs at the time). In her in-laws' home, she is forbidden her notebook. Needless to say, she will never become a great poet.

The story recalls Virginia Woolf's parable of Shakespeare's sister.[1] For our purposes, however, the feminist themes, although deeply important, are secondary. They are secondary for the simple reason that they are the product of something else – the way authors produce particular stories.

To consider this issue more carefully, we need to back up for a moment. Any study of literary narrative may be concerned with a number of different levels of generality. Thus a narrative theorist might be concerned with cross-cultural patterns – narrative universals that cover a wide range of works in different times and places. Alternatively, he or she might wish to focus on more culturally and historically specific patterns. More narrowly still, a theorist might be concerned with the works of one author or a single text.

Much previous work in cognitive theory of narrative has been concerned with general principles, cross-cultural or cultural. It has rarely focused on the particularity of authors or the particularization of individual texts. This is not to say that theorists have not analyzed unique literary works. They have done

xi

so extensively. Indeed, the main work of narrative literary analysts, cognitive or otherwise, commonly involves drawing on narratological theory to explicate particular texts. But that is a matter of *criticism*. To say that cognitive *theory* of narrative has tended to ignore particularity is to say something else. It is to say that there has been relatively little theorization of what makes narratives particular, how they come to be specified.

For example, *The Mind and Its Stories* and *Affective Narratology* set out general principles governing literary universals. They isolate broad patterns of genre structure. Moreover, the second book takes up those patterns to interpret individual works. Neither work, however, devotes much attention to the ways in which individual works come about. In keeping with this, they both discuss the structure of romantic tragicomedy. *Affective Narratology* examines a Sanskrit play, *Abhijñānaśākuntalam*, in terms of this structure, but it has little to say about how the author, Kālidāsa, came to produce that work. Put differently, both books isolate some cognitive structures that are very general – narrative prototypes. It is clear that those narrative prototypes play an important role in the production of many particular narratives. It is also clear, however, that specific narratives are not simply the reproduction of the prototypes.[2] For example, *Romeo and Juliet* is not the same as *Abhijñānaśākuntalam*, even though they are both romantic works. If the prototype is nearly the same across cultures, why are the individual works so different? Moreover, why are works by an individual author often recognizable as sharing features distinct from those that either occur cross-culturally or are found elsewhere in the author's historical period or culture?

Here we may return to Tagore. Readers of "Exercise-Book" are likely to feel that the story is very characteristic of its author. Moreover, it is different from other stories of familial separation, both those within Tagore's traditions and those from elsewhere. Even by placing Tagore's story in the context of his memoir, we begin to get some sense of what is going on in the particularization of this work. Tagore had emotionally important experiences of writing as a child. When he composed a story and reflected on his position as a writer, some of those memories were "primed."[3] In other words, some of those memories were partially activated but not necessarily made the object of focal attention. Primed memories help orient our self-conscious processes, pointing our thought and action in certain directions, affecting our inferences and judgments of likelihood, coloring our emotional responses, and so on. The memories in this case involved a child developing capacities as an author through practice with a notebook and through the encouragement of an older sibling.

We cannot precisely reconstruct what happened in Tagore's development of "Exercise-Book." We can, however, infer that there are two key elements in this development. One is the broad, cross-cultural prototype that provides the basic narrative organization of the story. This is the prototype of familial separation and reunion[4] – although here it is in its tragic version of familial separation and unfulfilled longing for reunion. This prototype is combined with the notebook and brother memories to produce the kernel of a story. The prototype subjects the autobiographical memory to a change. Now, the budding writer is separated from his brother. In Bengali families at Tagore's time, such familial separation would not occur normally for boys; however, it did occur normally for girls. Specifically, girls were married at a young age and, once married, were sent to live with their in-laws. This early separation was also a recurring concern in Tagore's stories. Thus this topic was regularly primed for Tagore and available for contributing to the formation of any new story.

When these elements are combined, the result is that Tagore must reimagine his own childhood experiences with a systematic alteration. The alteration is that the writer is now a girl rather than a boy. As he follows out his imagination of this single change, embedded in a familial separation narrative, he finds that the results are devastating. Uma (the girl in the story) will not grow up to be Tagore; society simply does not allow it.

Needless to say, these few observations do not constitute a general account of the generation of particular stories – far from it. They do, however, begin to suggest the rudiments of such an account. Most importantly, they point toward simulation as a central component of particularization. Simulation is our ordinary cognitive process of following out counterfactual or hypothetical trajectories of actions and events in imagination.[5] It is what allows us to get some idea of what it might be like to, say, ask the boss for a raise, before actually doing it. The operation of simulation in literature, first stressed by Oatley ("Why"; see also *Passionate* 171–181), has at least one highly significant consequence. Insofar as literary particularization is a function of simulation, it is continuous with our ordinary cognitive processes of counterfactual thinking. Imagining what Hamlet will do in his meeting with Ophelia is, then, directly comparable to imagining what one's boss will do when one asks for a raise. Thus it is open to the same descriptive precision and the same explanatory rigor. On the one hand, this means that our understanding of counterfactual simulation will contribute to our understanding of literary production. On the other hand, there is a wealth of detail available in literary works, an extensive elaboration or development of simulation that is rarely to be found in ordinary life. This suggests that literature is likely to provide a vast new body

of data to enrich and complicate our understanding of quotidian simulation as well. The focus of this book is on literature. In the course of investigating literature, however, it also develops an account of simulation that should be broadly applicable. In other words, the conclusions of the following analyses are only partially confined to authors of literary narratives.

At the same time, not all the conclusions in the following pages are generalizable. What authors do in making stories must be continuous with what we do in ordinary life. But there are differences. Smith's conjectures on the responses of his department head to some request simply do not call for the same plenary, communally shared attention as Shakespeare's *Hamlet*. Both are products of simulation and thus share some features, but they differ significantly as well.

Thus the following analyses, if successful, should contribute to our understanding of simulation generally. As such, they not only draw on, but are part of cognitive research programs in this area. At the same time, they should also contribute to our understanding of the distinctive sorts of simulation that guide literary creation. This too is part of cognitive research, but now in the area of verbal art. This is, presumably, just what a cognitive theory of narrative should do. On the one hand, it should build on and advance our understanding of the structure and operation of the human mind as developed in cognitive science. On the other hand, it should build on and advance our understanding of literary art as developed in narrative theory.

Chapter 1 takes up a fuller account of simulation. It explains simulation's basic principles and explores their evolutionary functions and mechanisms as well as the relationship between simulation and *theory of mind*. (Theory of mind is our capacity to understand other people's mental states – their beliefs, emotions, goals, and so on; see, for example, chapter 3 of Doherty.) The chapter also explores the precise manner in which simulative processes operate. To examine this topic more clearly, it introduces some ideas from neural network theories (or connectionism). These theories model mental processes on brain processes, using a limited architecture of units, connections, and activations that produce changing associative grids. The chapter goes on to argue that an underlying neural network may form the substrate of a system that is more perspicuously described in terms of general rules and representational contents (e.g., images). For example, a grammatical system for plural formation may be manifest in the human brain as sets of neurons, neuronal connections, and activation patterns in circuits of neurons. It may be clearest, however, to express plural formation in terms of a rule operating on a representation (e.g., governing when to add "s"). Recognizing the ultimate

intertranslatability of these approaches allows us to shift as needed between, roughly, associative and rule-based accounts of simulation.[6]

It is important to stress here that the distinction between neural-network and rule-based approaches is a matter of general explanatory architecture. It is not a matter of specific theories. The following analyses do not presuppose any particular connectionist or rule-based theories (or theories that combine connectionist networks and rules). Rather, they take up the two broad forms of cognitive architecture used for explanation. Indeed, the rule-based account does not even require that we think of authors as following rules (although, for ease of exposition, subsequent chapters will sometimes adopt this idiom). All that is needed is that we think of authors' activities as (roughly) conforming to rules or principles. That is why the two architectures are ultimately compatible. The underlying architecture is the neural network of the brain, but the psychological and social manifestations of that underlying network are regulated patterns best captured in principles.

Because the theory of simulation can be rather abstract, Chapter 1 goes on to develop and illustrate the central ideas of simulation in relation to literary works. Specifically, it presents within this framework detailed analyses of scenes from William Faulkner's *The Sound and the Fury* and Jane Austen's *Emma*. Faulkner's novel is one of the landmarks of interior monologue and thus a major work of psychologically oriented fiction. In this regard, it is particularly apt for study in this context. The choice of Austen may be less obvious. However, recent critical work (prominently that of Lisa Zunshine) has stressed the nuanced ways in which Austen explores people's understanding of one another's mental states – their theory-of-mind capacities. Thus Faulkner seeks to expose the interiority of characters' minds, whereas Austen examines how we understand – and misunderstand – that interiority when it is not exposed.

Chapter 1 examines story particularization through the more basic but potentially unwieldy architecture of neural networks. Chapter 2 turns to the level of rules and related higher-level cognitive structures, such as prototypes. It does this in order to explore what kinds of cognitive elements contribute to story simulation and what the process of such simulation might be. Specifically, this chapter outlines seven cross-culturally recurring story prototypes. It also presents a two-stage account of literary creation. In this account, one complex of rules governs the development of literary texts, specifically the alteration and specification of prior, more general structures (such as narrative prototypes). Another set of rules or processes governs an author's receptive evaluation of those texts – his or her judgment that the text

in question is or is not complete and final (i.e., that it does or does not require further development or revision).

Chapter 2 also distinguishes three types of simulated object – event, character, and scene. The narrative prototypes give the outline of events with some very limited constraints on character and scene. Our simulative capacities, in contrast, seem particularly oriented toward imagining personalities and intentions, thus character. Part of the argument of this chapter, then, is that an author's particularization of a narrative prototype is to a considerable degree the result of simulating a character into the event sequence of the narrative prototype. We have already seen a case of this sort where Tagore appears to have imagined himself as a young writer into the family separation sequence. As this case illustrates, the process is a complex one. As soon as the conjunction of character simulation and prototype occurs, there are changes on both sides. In Tagore's case, the separation becomes a marriage and the character becomes a girl.

Again, readers of Tagore would be likely to see "Exercise-Book" as highly typical of him. By the account presented in Chapters 1 and 2, that is not surprising. One author's emotional and episodic memories – thus his or her propensities for character and scene simulations – are likely to differ from those of other authors. Moreover, the precise clusters of development and evaluation rules will not be the same from author to author. Indeed, even the author's narrative prototypes are likely to vary somewhat from the prototypes of other authors. In other words, there is variation across authors' narrative idiolects – the complexes of memories, rules, prototypes, and other cognitive and affective components that contribute to their creation of individual stories. Thus it becomes crucial to examine simulation in relation to the cognitive operations of particular authors. That is the project of Chapters 3 and 4. In keeping with this change, both chapters signal a shift from the more abstract discussions of Chapters 1 and 2 to the more concrete analyses of the remaining chapters. Although all chapters involve both theoretical development and practical, illustrative analysis, the relative proportion changes at this point.

Chapter 3 sets out to isolate some of the distinctive features of Shakespeare's narrative idiolect insofar as this bears on his production of heroic works. This chapter focuses first of all on story structure, as did Chapters 1 and 2; however, it also pays attention to the emotional and thematic issues that are an important part of narrative production. (The two main purposes of storytelling cross-culturally are to engage an audience emotionally and to communicate some general normative points, usually political or ethical.)[7] Indeed, part of this chapter's argument is that Shakespeare's narrative idiolect exhibits significant continuity in thematic and emotional concerns. Specifically,

Shakespeare's heroic plays repeatedly oppose the usual political tendencies of the genre, systematically cultivating ambivalence. That general orientation, then, has consequences for the more particular development and evaluation principles that give rise to his stories.

Chapter 3 focuses on Shakespeare primarily because of his unparalleled stature in world literature as well as his general familiarity. The stature suggests that his work should reveal successful patterns in the production of verbal art. Familiarity also makes his work particularly suitable for introducing and illustrating the basic ideas contained within the remainder of the book. In addition, Shakespeare has been perhaps the central figure in the recent development of historicism. Many readers are likely to see the cognitive approach of this volume as opposed most sharply to new historicism and cultural studies. In fact, this is a false opposition. Ultimately, cognitive and historical/cultural studies require one another. Nonetheless, given the current critical context, it seems particularly apt to introduce cognitive principles in relation to the primary interpretive target of new historicism, thereby conveying the relevance of the former even for texts closely associated with the latter.

Chapter 3 necessarily treats differences among Shakespeare's heroic narratives, even while concentrating on the shared idiolectal features. However, it does not isolate any particular pattern in those differences. Chapter 4 turns to the issue of how one may speak of an author's narrative skills as developing or advancing across a canon of works, as opposed to degenerating or falling into repetition. To explore this topic, the chapter further articulates the idea of a rule in narrative idiolect. Specifically, it presents an account of cognitive operations according to which rules may repeatedly be revised to include variables or parameters. These parameters interact with one another such that the setting of one parameter leads to cascades of changes in other parameters. For example, a character may be understood as a cluster of principles. Changing the sex parameter in one of those principles is likely to have consequences in a range of others – as well as associated principles bearing on, for example, events and scenes. Returning again to Tagore's story, we can see the shift from Tagore to Uma as involving, first of all, a shift in a character parameter from male to female. This leads to a number of other shifts in parameters. For example, the brother stops being unequivocally supportive of the sibling's literary creativity. This is not simply accidental; it is connected with the differing norms that bear on the activities of boys and girls. Moreover, the places where the child spends his or her time change, as does his or her general authority in the household.

To examine these issues more thoroughly, Chapter 4 takes up Jean Racine's secular tragedies, which are all partially romantic. As explained in the

chapter, Racine is a particularly apt figure for this examination as his writings are strictly segregated from one another in a known chronological sequence. This allows us to isolate changes in his narrative particularization across time. In addition, Racine's work has been seminal for the development of modern critical theory. This is shown not only by Roland Barthes's early work, *On Racine*, but also by Barthes's explanation of the critical importance of "Racine's contemporaneity." Specifically, Barthes explains that at the time of his book, Racine's "work [had] been involved in every critical effort of some importance made in France." He lists "sociological … psychoanalytical … biographical … [and] phenomenological" schools, concluding that Racine is the only "French author … to have made all the new languages of the century converge upon himself" (viii).

The sorts of simulation examined up to this point are largely a matter of character simulation within a prototypical event sequence. This is appropriate, as it seems to be the key manner in which stories are particularized. There are, however, other elements that enter into the specification of plots. The importance of theme suggests that one such element is argument. Authors are sometimes guided in their imagination of a story by what are in effect patterns of reasoning. Indeed, there are suggestions of this in all the authors considered thus far, although it may not be the most prominent feature in any of these cases. Such reasoning is most often integrated with narrative prototypes and character simulations. Nonetheless, it is important to consider particularization through argument separately.

Authors' use of ideas in simulation is not confined to abstract sequences of reasoning in the service of thematic goals. Authors often take up another sort of ideational process in simulation as well, and not necessarily for thematic ends. This is modeling. A model may come in the form of a precursor text (as in Joyce's use of *The Odyssey* in *Ulysses*). It may be found in characters and events shared by a tradition, as when an Indian author imagines two human lovers on the basis of the paradigmatic spiritual lovers, Kṛṣṇa and Rādhā. It may be found in a broader metaphor, as when an author makes a character blind to express some lack of insight. In each case, what we have is not a logical argument but rather some sort of analogy – an analogy that may or may not be used self-consciously. A model may contribute to the communication of themes (in the manner of an argument). However, it may also enhance the emotional effects of the work, orient the simulation of character, or guide the development of events.

Chapter 5 takes up these two modes of particularization. It begins by discussing how, cognitively, the simulation of a work may be guided by an argument. It then examines Bertolt Brecht's "teaching piece," *Die Maßnahme* (*The*

Measures Taken), "one of Brecht's greatest achievements" (43), "his first real masterpiece," according to Martin Esslin. The chapter considers what Brecht's argument is in this play and how it shaped his particularization of the story – or, rather, how initial ideas interacted with narrative prototypes and other factors to produce a simultaneous particularization of the story and of the argument. Brecht is an obvious choice for this analysis given his teaching project in this and other plays. Moreover, *Die Maßnahme* is framed as a criminal investigation, and its embedded narrative is sacrificial in structure. This adds two other genres to those examined in the course of this book, following the heroic (in Shakespeare) and romantic (in Racine) – thus four of the seven recurring cross-cultural prototypes (as isolated in *Affective Narratology*).[8]

Chapter 5 goes on to consider the operation of cognitive models (or metaphors) in guiding simulation. It first treats this process in general and then examines Kafka's *Die Verwandlung* (*The Metamorphosis*), also a sacrificial structure. It may seem that Kafka's novella is entirely straightforward in this regard, drawing on a single, simple model – Gregor is an insect. In fact, this case is interesting precisely because there are several levels of modeling that interact with one another in the process of simulation to produce the final story. An awareness of these models helps clarify the complex operation of modeling in the specification of stories. At the same time, it helps reveal nuances of this particular work that may not otherwise be apparent.

The remaining chapters continue the focus on individual narratives. Now, however, they turn to discourse. The standard distinction between story and discourse is between what is told in the narrative (the story) and how it is told (the discourse). The story is the events, characters, and scenes as they may be supposed to exist "really" in the narrative. Discourse is divided into two components. One concerns the facts of the story world – what is reported, when it is reported, and how it is reported. This is called *plot*. Plot varies the selection, construal, and organization of story information. For example, in a detective narrative, readers commonly learn about the crime first and only learn about the motive for the crime at the end. Thus the plot begins with the crime and ends with the motive. The story, however, begins with the motive and proceeds from there to the crime. The second component of discourse does not concern what the reader is told, but (roughly) who does the telling. For example, in the case of a criminal trial, different witnesses may be tellers of more or less the same story events. Moreover, their narratives will be embedded in an encompassing narrative – for example, a juror's recollection of the trial. This part of discourse is called *narration*.

Chapter 6 and the Afterword concern the ways in which authors particularize narratives through emplotment (the formation of plot) and narrative

voice. Discourse is most widely treated in prose fiction. Emplotment, however, is clearly an important feature of drama as well. In order to consider that somewhat underexamined mode, Chapter 6 takes up a play, *Hamlet*. It considers the ways in which the selection and ordering of story information serve to particularize the work and how this particularization serves emotional and thematic purposes. It also considers the ways in which emplotment manifests recurring principles of an author's narrative idiolect.

The other aspect of discourse, narration, is highly diverse and complex. Theorists devote entire books to a single aspect of narration, and it is not possible to give anything approaching a full treatment of this topic in a book devoted primarily to the story component of narrative. The Afterword takes up some selected topics in the particularization of narration, stressing the simulation of narrative voice. A central argument is that narrational simulation involves exactly the same processes as story simulation but with different initial prototypes. A prototypical scene of storytelling is not the same as a prototypical romance, but the imaginative processes are the same in both cases.

The simulation of narrative voice is, in fact, general and applies to academic discourse as much as it applies to fiction. The main difference is that the simulation of a narrative voice in academic discourse is commonly concealed. In acknowledgment of this, the Afterword takes up the issue of narration not only at the level of argument but also at the level of narration itself, presenting the analysis through a series of different voices. These voices are drawn in part from the authors whose works are the topics of this chapter – William Faulkner and Italo Calvino. Moreover, the specific issue addressed in relation to these authors is how one might treat the artificial quality of narration. Telling a story commonly involves, for example, concealing key information until the end. It is, as such, a highly contrived process. Authors may deal with that contrivance in different ways. The options of naturalizing or explaining the artifice versus flaunting it are neatly exemplified by Faulkner's *Light in August* and Calvino's *Se una Notte d'Inverno un Viaggiatore.*

Thus, in addition to presenting a general account of narrative simulation, the different chapters of this book examine patterns found across an author's works or within an individual narrative in order to isolate the principles that generate particular stories. In other words, as the title of the book indicates, they explore some aspects of how authors' minds work. These chapters repeatedly stress that the principles at issue are implicit, thus they are not something that authors are able to introspect. One result of this is that authors' testimonials are largely irrelevant to the following analyses. It is crucial to emphasize this at the outset. To be concerned with how authors' minds make stories

is not to be concerned with how authors imagine that their minds work or how they explain creativity when pressed to answer a question in an interview. An author might believe that he or she has been inspired by God (as may have been the case with Harriet Beecher Stowe and *Uncle Tom's Cabin* [see Kazan x]). This hardly means that a cognitive approach to literary particularization – in Stowe's case, often the particularization of a familial separation and reunion prototype – should make reference to divine providence.

The point is not by any means peculiar to authors; it applies across the board. Consider, for example, ordinary speech. As we will discuss in Chapter 1, fluent adult English speakers are very bad at stating the principles that govern English grammar. In their speech, they conform to the principles. But if they seek to formulate those principles on the basis of introspection, they fail. If anything, this failure is more likely in the case of complex activities, such as the writing of plays and novels. In keeping with this, Bargh points out that there is "a deep and fundamental dissociation between conscious awareness and the mental processes responsible for one's behavior; many of the wellsprings of behavior appear to be opaque to conscious access" (560). Indeed, the point even applies to goal pursuit. Bargh cites research showing "that social and interpersonal goals can … be activated automatically through external means.… The individual then pursues that goal in the subsequent situation, but without consciously intending to or being aware of doing so" (562).

Such failures of introspection are what make cognitive analysis necessary. If the principles of grammar or other human activities were intuitively obvious, there would be no need for cognitive research. Linguistics, psychology, literary theory, and a host of other disciplines would be eliminated as their findings would be self-evident without research. Unfortunately, that is not the case. The following chapters, then, are an attempt to respond to this human deficit in intuitive self-understanding with respect to literary particularization – that is, with respect to how authors' minds make stories.