

THEORY OF MIND & THE SCIENCE OF STORY

Buckets of rain
Buckets of tears
Got all them buckets coming out of my ears
Buckets of moonbeams in my hand
You got all the love honey baby
I can stand.

- Bob Dylan, *Buckets of Rain*

I once loved two women simultaneously—two women with whom I shared in intimate, sexy relationships of varying levels of intimacy and sexiness. Rarely was one far from my mind when I was with the other, and though I was constantly confronted with making a final, hard-line choice between the two, I didn't do well at all. Just when I'd think I could see clearly, that the choice was obvious as the day, one would reveal herself to be so self-serving, or the other would do something so perplexing, that my bifurcated heart could not truly settle. In the end, what saved me was that neither of these women was real. They were Frieda Winter, the obscenely wealthy New York sophisticate, and Jeanne Green, the modest editor with a true eye and truer heart, and they were the long-term love interests of the brilliant but fictional writer Arthur Youngblood Hawke.

Around the turn of the century, my sister gave me Herman Wouk's *Youngblood Hawke* after hearing a writer friend tell his conversion story: he was a successful but not wholly satisfied engineer; he read and devoured Wouk's 700-plus-paged tome, stopped his career on a dime, and began to write. It's that kind of book. Emotion, revelation, physical pain—all are seen and felt viscerally, as Arthur experiences them. Arthur Hawke is a definitively large character; part of Wouk's magic is the way in which Arthur, who we

THEORY OF MIND & THE SCIENCE OF STORY

follow the full course of his writing career, becomes more frailly human and more preternaturally brilliant as his story unfolds. He's a wonderful paradox, but at several points he feels like something other, an ET—until Jeanne or Frieda re-enters the scene. They're rendered with such beautiful detail, as different from the other as black and white; each with that touch of the ethereal, making them easy and frustrating to love.

The character of Arthur is hard to contain; he's unruly and kind, very stubborn and very much of a genius, prone to extremes. In my personal literary lexicon, Arthur holds the position of the protagonist who has afforded me the most insight into supporting characters by way of his own impenetrability. These two women, Jeanne and Frieda, they couldn't figure him out, either, and so I latched onto that commonality on their end. If you're scoring at home, so far we have me

- a) loving both Jeanne and Frieda because Arthur loves the beautiful and perplexing both of them,
- b) loving them because his love is often his most accessible trait, heightening my connection to him, and
- c) loving them because they don't understand Arthur, and I don't, either, so I feel their pain and subsequently feel I understand them.

The duality of my love for these two very different women was super confusing, as I read along. They represent the light side and the dark side, for Arthur Hawke, but neither properly possesses his heart, and, as the story moves along, the contrasts become less stark. I could list for you the features

and facets of these two women—an unequalled intensity of gaze, a tantalizing verbal acuity—but you don't need me to.

please hear me : theory of mind explained

If you read novels, you've been in love. And if you've been in love, you know it is work. Knowing the mind of the one you're with is *work*, and holding all the micro-levels of two women's needs and wants, let alone keeping them separate, drove Arthur Hawke to the brink, and it was just plain beyond me. But love and attraction aside, the mind of Jeanne Green held a curiosity that was a cut above, and made her the right choice. Frieda and Jeanne were both avid intellectuals, socially perceptive, and bright as all get out, so it wasn't that Frieda was not curious about the world. Rather, she was not curious enough Arthur—and, working by the transitive property of readership, not curious enough about *me*.

Theory of mind is most simply defined as the ability to know the intentions of another person. Within that framework, the two central components are the ability to know another's rewards and punishments *a priori*; and the ability to discern that these rewards and punishments are different from yours, but still valid in their world.

John Gottman is a storied University of Washington psychology researcher. He began his academic career as an applied mathematician, and his affinity for numbers have served him well—he's much more dexterous with his data than are most social scientists. Gottman's Love Lab has been tracking

THEORY OF MIND & THE SCIENCE OF STORY

couples for more than 20 years, and has developed a discreet scoring system for a married couple's interaction. They've gotten pretty good at it; Gottman can observe a couple for three minutes, and tell you with more than 90% accuracy if they'll be divorced three years later. Where I live, the Seattle area, more than 50% of married couples who complete a nine-month period of couple's therapy have filed for divorce thirty days later. In any standard model, couple's therapy goes out of business—it's a crapshoot. But Gottman's lab needs three minutes; and in observing a couple for ninety minutes, their success rate goes up to better than 96%.

When Gottman went looking for independent variables—determining factors in the success or failure of a marriage—they found one that superseded all others by a wide margin. Shocking to no one, this independent variable belongs to the woman, and it is: do I feel heard.

Not, interestingly, does he understand me, but rather, do I *feel* understood.

After I was done with all the analysis, a while passed before I was able to pin down the definitive problem in the Jeanne Green/Frieda Winter miasma, and once I did pin it down, I sort of wished I hadn't, because the gender roles were essentially switched. For us (translation: Arthur plus the vicariously loving me), with Frieda Winter, the breaking point was curiosity because she wasn't curious about a specific part of us: she was not curious about our mind. We did not feel heard.

theory of mind : having it makes us human

THEORY OF MIND & THE SCIENCE OF STORY

At the opening of his book *Stumbling on Happiness*, Harvard psychologist Daniel Gilbert writes,

Few people realize that psychologists ... take a vow, promising that at some point in their professional lives they will publish a book, a chapter, or at least an article that contains this sentence: 'The human being is the only animal that ...' We are allowed to finish the sentence any way we like, but it has to start with those eight words. Most of us wait until relatively late in our careers to fulfill this solemn obligation because we know that ... the worse we do, the better we will be remembered. For instance, those psychologists who finished The Sentence with "can use language" were particularly well remembered when chimpanzees were taught to communicate with hand signs. ... So it is for good reason that most psychologists put off completing The Sentence for as long as they can, hoping that if they wait long enough, they just might die in time to avoid being publicly humiliated by a monkey. I have never before written The Sentence, but I'd like to do so now, with you as my witness: *The human being is the only animal that thinks about the future...* We think about the future in a way that no other animal can, does, or ever has, and this simple, ubiquitous, ordinary act is a defining feature of our humanity.

Gilbert's book is chock-full of fun insights about the makeup of the human mind, and, despite the *happiness* of the title, his version of The Sentence is the work's central conceit. There are two ways in which it's relevant here.

First, I'm going to happily submit to you another rendition of The Sentence. *Our having theory of mind is what makes us human.* (A qualifier is needed here, which is why this version is not as strong as an opener for a book. In a few cases, animals display a rudimentary theory of mind, most notably

chimpanzees.) To be Ziploc-tight about it, we need to say: our having *complex* theory of mind is what makes us human.

Second, we'll come back to Gilbert's thesis in a bit, and we'll posit that the human ability to think about the future is actually a tributary of theory of mind—our ability to think about each other's minds.

the king is dead, so is the queen : story vs. plot

Some years ago, psychologists framed the levels of social cognition as *intentionality*. Being aware of your own thoughts constitutes a first order of intentionality. Awareness of someone else's thoughts constitutes a second order, and this is the baseline for theory of mind. An example of a third order might be, Huntsman *suspects* (1) that the reader *feels* (2) that the author merely *wanted* (3) to have the essay wrapped up before skipping town for Vegas. What the orders of intentionality do is allow us, by quantifying the degrees of separation in a given instance of theory of mind, to calculate the level of sophistication inherent in that instance. Within literature, this system of orders is an efficient way to look at theory of mind, and see how it's vital in granting layers of motivation, and meaning, to a story.

In his excellent book *The Human Story*, behaviorist Robin Dunbar uses Shakespeare's *Othello* to explain the orders of theory of mind.

Iago had to *intend* (1) that Othello would *believe* (2) that Desdemona *wanted* (3) to love another before it was worth his while saying anything at all to the Moor... It was Iago's instinctive ability to see how Othello would interpret the information about

THEORY OF MIND & THE SCIENCE OF STORY

Desdemona's intentions—and the audience's ability to foresee the awful inevitability of that interpretation and *anticipate* its consequences—that makes the play work. Had Iago not been able to engage in these mental gymnastics, he would not have been able to feed Othello a bunch of lies. Othello would have remained in sublime ignorance of Desdemona's presumed behaviour and he would never have killed her; nor, in his anguish at realizing how he had misread the situation, would he have gone on to kill himself. In which case, the story would have lost most of its emotional force. Without third order intentionality, Iago could not have done what he did. Without fourth order intentionality, we the audience would not have been able to figure out the big story line. And without fifth order, Shakespeare could not have put the whole thing together and manipulated our minds the way he so miraculously intended to do.

Without fifth-order intentionality, Shakespeare would've been little more than the proverbial chimpanzee at the typewriter, hitting keys at random. Without theory of mind—indeed, without higher orders of it—literature would be impossible. And it probably wouldn't be worth having a conversation with anyone.

How far can the human mind go—how many orders of intentionality can we do? Six, at most. Five is a not inconsiderable feat—in a Dunbar study, participants were read a sequence of events, and were found capable of reciting back generally four and no more than six layers of intentionality. These same participants, however, could recount nine to eleven layers of linear action, eg, Phil *woke (1)* up and then *fed (2)* the dog but the dog *ran (3)* away and so Phil *went (4)* to the store and *bought (5)* food for breakfast, which he then *cooked (6)* and then *ate (7)*. Humans have a generally astute memory for linear events;

THEORY OF MIND & THE SCIENCE OF STORY

meanwhile, we top out at six orders of intentionality, absolutely. And if you can get to five, you're doing good work.

The linked sequence of events involving Phil, his dog, and his breakfast is a decent example of what, in his seminal book *Aspects of a Novel*, EM Forster defines as a story. He writes, “‘The king died and then the queen died’ is a story. ‘The king died; and then the queen died of grief’ is a plot.” We use *story* to describe a number of things, so it's useful to have Forster's definition offer such tight parameters. The magic of his differentiation, however, comes to two small words: *of grief*. Those are all that's needed to grant a psychological insight to the queen, and suddenly our ears have perked: she loved him; she loved him so much it killed her, how tragic. Forster's definition of plot, then, can be described as a story wherein action is motivated by intention.

For a work of fiction to be properly moving, let alone to aspire to art, it must be a narrative that has complex levels of psychological motivation within it. Theory of mind is what we want to read, and we want to watch it progress. We want to read characters' respective theories of mind changing about each other.

The actors by their presence always convince me, to my horror, that most of what I've written about them until now is false. It is false because I write about them with steadfast love (even now, while I write it down, this, too, becomes false) but varying ability, and this varying ability does not hit off the real actors loudly and correctly but loses itself dully in this love that will never be satisfied with the ability and therefore thinks it is protecting the actors by preventing this ability from exercising itself.

- JD Salinger, *Seymour: An Introduction*

THEORY OF MIND & THE SCIENCE OF STORY

No narrative that tells us the facts of a man's life in a man's own words can be uninteresting.

- Mark Twain

an introduction : narrative defined

If we confine ourselves to the either/or framework of story vs. plot, then our conversation is bound to ignore that more ethereal feature of literature— narrative. Narrative is tricky; not, as it turns out, overwhelmingly complicated, but tricky. It lends itself to multiple usages, none of which are clearly, commonly defined. Meanwhile, plot, even properly understood to contain psychological motivation, is an outlier, a concept—a story has a plot. A story can have a narrative, or it can be rightly said to *be* a narrative. Between the two terms, narrative is the more robust descriptor; something a story is rather than has. Take JD Salinger's *Seymour: An Introduction*. Made up of a series of letters written years after Seymour's death by Buddy, his younger brother, *Seymour* is a stream-of-consciousness collection of reminiscences and propositions, and simply doesn't offer itself to a linear progression in Forster's plot mold. Charting *Seymour died; and then, years later, his brother wrote down a bunch of loosely associated things about him and their family*, is not going to earn you a gold star from teacher.

Here's another one. A theory of mind-inclusive plot summary of Saul Bellow's *Herzog* might run something like this: Moses Herzog wants to escape the city and goes to Martha's vineyard but comes back right away; Moses calls his lawyer and goes to the courthouse; Moses thinks about murdering his ex-wife but doesn't; Moses spends the night with a friend, gets in a car accident,

THEORY OF MIND & THE SCIENCE OF STORY

gets arrested, gets out, and goes to the Berkshires; Moses enjoys the pleasant weather in the country and feels content. That may seem a fair-sized collection of events, but I shouldn't think it jumps out at you as a solid 400 pages of action. To even approximately talk about the book's undertaking, you need to say what it's about: *Herzog* is a narrative about a man trying to find himself. Narrative, then, while a broadly applied and generally slippery term, is useful in describing both what a story is and what it does. Let's define it simply:

Narrative is theory of mind changing per unit of time.

In any case experience folds upon itself, refers backwards and forwards to itself through the referents of hope and fear; and, by the use of metaphor, which is at the origin of language, it is continually comparing like with unlike, what is small with what is large, what is near with what is distant.

- John Berger

don't be such a monkey : theory of mind—the roots of metaphor

In Robin Dunbar's *The Human Story*, he relates the story of a client of his, a British woman with an autistic teenaged son. Autism disorder occurs across a broad spectrum, and this boy was towards the high-functioning end. One day they're on their way to the grocery; she tells son, "Pull the door behind you," and goes out to the car. (In British English, *pull the door* is the equivalent of the American *shut the door*.) A long moment later, out comes the son, carrying the door under his arm. What's going on here?

There are several mystifying things about autism. Beyond not having a medical treatment for it, for instance, we don't know why it has such a higher occurrence rate in boys than in girls. But by studying autistics—behaviorally,

THEORY OF MIND & THE SCIENCE OF STORY

and by using fMRI imaging technology to take pictures of the autistic brain's activity during mental tasks—scientists have been able to learn a lot. For our purpose, the thing to know is that autistics are incapable of perceiving the world, of taking spoken meaning, as anything other than literal. You want I should pull the door behind me? Well, okay. Just don't accuse me of jumping the gun when there's clearly no gun in the room.

Among psychologists in the 1950s and early 60s there was much excitement over the idea of teaching chimpanzees to speak. The endeavor hinged on the question of whether human children learn language instinctively, or because growing up in a community offers them a learning opportunity they cannot avoid. Families of American psychologists raised chimps in their homes, in some cases alongside their own offspring, where everything done with the child was done with the chimp. And the efforts were a success, in the barest sense: chimps did learn to speak a few human words. Beyond that, they were a full-blown fiasco; a failure of comic proportions. Chimpanzees grow much faster than humans, and provided atrocious models of behavior for the experimenters' children—human kids are superbly specialized imitating machines, and will readily imitate whatever is on hand, particularly if it's a few shades on the side of naughty. These experiments abandoned, and with good reason never repeated. But Dunbar recounts two more interesting reasons why the chimp-in-the-family endeavors failed.

First, the best a chimpanzee could do was whisper sounds it was imitating—chimps simply lack the vocal apparatus to speak. Their larynx is set

THEORY OF MIND & THE SCIENCE OF STORY

high in the back of the throat, just below the tongue, whereas human's is set low in throat. (Human babies are born with a high larynx, which drops as they grow into the language-learning period, and this is just as well, for a high larynx lets infants swallow and breathe at same time; they can nurse without stopping to breathe.) Humans need a low larynx because it enlarges the resonance chamber in the throat and mouth—a necessary component in producing the range of sounds used in language.

The other reason is that chimpanzees lack theory of mind. Noam Chomsky put linguistics on map as official science more than fifty years ago. In that time, linguists have focused on grammar, and how, as a universal framework, it allows us to transmit info by encoding it in a sound stream of speech. Hard as this may be, it is not most difficult component of speech: the real intellectual work lies in anticipating what the listener is going to understand (or *not* understand). The task of talking to a person, and, as you look at them, instantaneously tailoring the content and delivery your speech—guessing what they will understand, and accept, enjoy—is so profoundly complex, and the ease with which we do it so automatic, that a proper analogy is difficult to create. So, rather than analogize, let's empathize: imagine speaking like a fairly high-functioning autistic. Imagine being limited to uttering right-as-rain grammatical sentences that describe the world as it appears. *The shade of green on the grass right here is different from the shade of green over there.* It's a world where theory of mind is absent, and so it's utterly lacking in passion, humor, higher understanding; and lacking metaphor, or secondary meaning.

THEORY OF MIND & THE SCIENCE OF STORY

We tell jokes, we use metaphor so liberally that almost nothing we say has its literal meaning... We speak in riddles and circumlocutions. We go to great efforts, it seems, to avoid saying *exactly* what we mean in plain simple English. For all the difference it would make, we might just as well be speaking French, or Polish or Chinese.

– Robin Dunbar, *The Human Story*

Looked at through a theory of mind lens, metaphor can be understood as a sort of symbolic reasoning, requiring not just theory of mind but higher orders of theory of mind. From an evolutionary perspective, the task of understanding other human minds in a social context was sufficiently complex that the human brain developed frontal lobes large enough for language, and a cortex—a wafer-thin layer of eight or so cells that covers the rest of the brain—with the neuron power necessary to understand symbolism and secondary meaning. We needed this capacity, because as the African jungle gave way to the savannah, early humans had to mobilize, and their survival was inexorably linked to being able to work in a large group. And, if you and I are to have a conversation as we work and walk and rest on the savannah, very soon we're going to need to discuss and agree about someone else, a 3rd party—and that, in theory of mind terms, requires 3rd order of intentionality. We're not just reading the mind of each other, but hypothesizing the intent and contents of a third mind, on which we must agree.

If we're to avoid a sort of paralysis of failed understanding, we must engage in a sort of mind-reading. And we don't emerge from the womb knowing how to do it—it's an acquired skill. People's minds are very complex to understand and guess at; this being the case, let's start with objects.

THEORY OF MIND & THE SCIENCE OF STORY

The ability to recognize that an object can have multiple functions—eg, that a bag labeled *chips* could hold onions or pens or whatever else, instead of chips—is called Dual Representational Theory, and was first observed and named by Judy DeLoache, a wonderful psychologist at the University of Virginia. She says, “To use a symbolic object such as a model, map, or picture, one must achieve dual representation. That is, one must mentally represent both the symbol itself and its relation to its referent.” Is this not truly fascinating? A thing we adults do so readily it seems pre-programmed—say, looking at a map and recognizing that it refers to a much larger physical place—is in fact a task that needs cognitive function of such a robust order that the brain of a child who can speak in complete sentences has not yet built the network needed to do it.

My nephew Kai (three years old) is a brilliant storyteller and expert builder of Lego boats and cars and the garages that hold them; each time I have the pleasure of playing with him, I am freshly amazed at some new aspect of the world his perspective has grown big enough to entertain. All the same, I would never undertake to explain to him the allegorical meaning of a story we’d just read or watched together. Of course not—I know intuitively that it’d be a waste of both our time. What my intuition does not tell me is that I could just as soon succeed in explaining to him that Lord Aragorn is a Jesus figure as I could fish out my star chart, go with Kai out to the backyard, and get him to see that the stars in the sky above correspond to the white dots on the page. Though very different in their sophistication, both ideas are equally impossible for Kai to recognize because both involve secondary meaning, for which his

THEORY OF MIND & THE SCIENCE OF STORY

brain has not developed sufficiently to comprehend. The human's is the only mammal brain that's not yet big enough at birth to perform all the tasks survival requires of it, and the primary reason for this is simple: if it grew any larger prior to birth, passage through the birth canal would not be possible without fatally damaging the mother; as such, humans are uniquely positioned as the most intelligent animal as well as the one incapable of surviving on its own for years longer than any other.

As with star charts, so with minds: because a three-year-old human cannot yet grasp that someone can have a belief different from his own, yet a four- or five-year-old does, we can pin this down as the window when theory of mind is beginning to come online.

If such a high percentages of what we say is not literal, but instead requires an understanding of metaphor (Dunbar puts the proportion of our verbal communication that has a strictly social function, rather than a logistical Xs and Os nature, at an obscenely high sixty percent) which itself is brought to you by theory of mind, then Daniel Gilbert's stance that thinking about the future is a defining feature of our humanness is actually dependent on theory of mind. For metaphor is what enables us to think about time: things in the past happen *behind* us; events in the future happen *ahead* of us. We are able to think about time by making a time *line*—grafting a spatial metaphor onto our concept of time. I believe Gilbert is correct—thinking about the future is a defining aspect of humanness. But, working in reverse, so then too is metaphor, and theory of mind before that. From an evolutionary theory perspective,

THEORY OF MIND & THE SCIENCE OF STORY

At the meta-level, literature—and, really, any form of art—must also *require* theory of mind. Not just so we can understand a character's sarcasm, but so we can extract levels meaning from satire, or allegory, or any of the bigger, grander forms of metaphor.

There's an old joke about an elderly woman who immigrates to the United States from Eastern Europe. She's at Ellis Island, and the immigration officer is giving her a required oral test, steamrolling through the many questions. *Would you support the overthrow of the United States government by force or violence*, he asks, and she thinks about it and says—*violence*.

ivan unhinged : we are very, very good at theory of mind

Some years back, a writer-friend and I decided to co-write a story by email. We played around with the rules of play a few times; what we found worked best was trading paragraph for paragraph. Other than us both being surprised at how easily we matched voice—picking up each other's invented quirks with hardly an overt effort to do so—the major takeaway from the experience was a significantly increased understanding of what makes a paragraph count, what makes it feel whole: it has nothing to do with the length of said paragraph. What makes a paragraph whole is that it has *action*: sense of motion. Expanding that out, you could describe a sense of motion as a vital property of an effective narrative. Not super insightful on its face—even a two-bit pundit knows to pepper his sauce with active verbs—but, for me and my friend, it began a conversation and experimentation as to what *kind* of action we, as readers, find engrossing. On the one hand,

THEORY OF MIND & THE SCIENCE OF STORY

The king died then the queen died: Okay, sweet—that’s a narrative’s worth of action all by itself. But,

The king died, the queen died of grief ... oh, did she? How tragic, how awful for her; tell me *everything*.

The second has the motion of the first, plus an active resonance point, and it’s concentrated in two syllables: of grief. Two words amended to the end of a sentence, but suddenly there’s a resonance point we can empathize with—there’s a point of entry into the mind of the queen. We know this from before, via Forster—that theory of mind turns story into plot. But now we get to see why, in the realm of narrative art, this is a useful thing to understand.

Theory of mind is the connective tissue, and not merely in terms of providing (according to evolutionary theory) the neural framework that allows humans to think abstractly. It also allows us, as readers, to maintain emotional connection with a story or narrative. It doesn’t offer higher meaning on a plate, but it does what must precede that: it makes us care what happens to the characters. And it’s around that insight and empathy that greater meaning can take hold and begin to resonate with a given reader.

Or a given viewer. Because a common point of experience is worth pages and pages of anecdotal analogy from me, on the next page is an image of *Ivan the terrible and his son Ivan*, by Ilya Repin, a 19th- and early 20th-century Russian painter. Give it just a moment before you let more of my words crowd the field.



(an un-cropped image of the painting can be found here:
http://upload.wikimedia.org/wikipedia/commons/6/6c/REPIN_Ivan_Terrible&Ivan.jpg)

This work is more commonly known as *Ivan the Terrible killing his son*, and, indeed, that's what happened. His son's bride-to-be became pregnant, and

THEORY OF MIND & THE SCIENCE OF STORY

Ivan ordered a brutal abortion. When Ivan Jr. came to see him, enraged, they argued, Ivan the Terrible grabbed an iron curtain rod and struck his son in the head, fatally wounding him. Everything in the painting flows to those eyes, that partly concealed face. We don't even see his mouth, and we needn't: his eyes, brow, dented cheeks, and even his hands offer a traitor's condolence. We look at it and are able to jump right into a terrible storm of emotion, exquisitely rendered by Repin, and translated by us in a blink, thanks to our minds' ability to instantly extract a layered impression of a foreign human mind from a brilliant collage of discreet details. It's one of my favorite paintings.

Every time I look at *Ivan killing his son*—even now, it's still kind of hard to glance—I am taken by how much emotion is on his face, and I say to myself the same line.

I can't believe how f*cking good we are at this.

The task of watching another person, their speech, behavior, movements, and discerning their mood may seem like a matter of concentration—everything serves as a cue, so how big a deal is it, right? You talk to a family member for a decent length of time, you watch them, and even if they don't speak specifically to their mood, they may as well have. But here's the thing: it needn't be a family member or lover for you to do this, or a friend; or someone you know at all. It may be a stranger. It may be you open the door to a room and glimpse the face of an terrible, relentless tyrant, and, in what seems like no time at all, you perceive multiple layers of his narrative, and are entranced by the range of terrifying emotions that have collapsed in on him.

THEORY OF MIND & THE SCIENCE OF STORY

Let's try one more painting from Ilya Repin—this one entitled *They did not expect him*.



Where are (you) going (with this)?

By now you might think I'm cheating. Paintings are not literature. But while I came to the idea of narrative having a basis in theory of mind by way of literature, a painting's narrative can be experienced much more efficiently than a work of fiction. This painting, even more dear to me than *Ivan killing his son*, is exploding with theory of mind.

The man entering the room has been in a Russian gulag for eight years and was long thought dead. He was a friend of the activist Repin, considered by everyone to have perished, and then one day he walks in the door. Unpack it with me:

- A) The maid – apprehensive; she could close the door, but she does not
- B) The mother – she stands, alarmed as though she has seen a ghost
- C) The wife – eyes wide, whole face delighted but disbelieving
- D) The girl – taking her cue from the grownups, suddenly shy
- E) The boy – only he shows open pleasure

One thing to note is how this takes us back again to Daniel Gilbert's thoughts about the future. He says, "to see is to experience the world as it is, to remember is to experience the world as it was, but to imagine—ah, to *imagine* is to experience the world as it isn't and has never been, but as it might be." That ability to *imagine* extends not just to our ability to create art and imbue it with narrative, but also our ability to receive it. Try interpreting a scene like this without reference to human consciousness, to what *we* know of human feelings, and the depth, the human depth, completely disappears.

THEORY OF MIND & THE SCIENCE OF STORY

Look at that boy's face again – he is overjoyed. But also he is ... vindicated. *I told them*, his eyes are saying. *I said he'd come back. They wouldn't believe me but I **knew it**.*

Based on the title of a painting by Gauguin, *The Inner Eye's* final chapter is entitled *Where are we going?* This paper somewhat avoids that most-meta level of questioning, because, while it covers two very big and hard-to-define things—the human mind, narrative—the concern here has been how the first is wrapped up in our experience of the second. And I want to tell you that an understanding of theory of mind has fundamentally shifted the way I encounter literature, the degree to which I engage a narrative. It's not a matter of depth so much as access—applied to narrative, theory of mind is a framework for looking at a narrative, a set of parameters; and it offers an elegant means of getting at the mind of a character. It's a phenomenally complex undertaking, the reading of and guessing at the minds of others, even ones right there on the page. *Elegant* is the right word because we're so good at it, built for it, we read minds all day long, whether we're readers or not—each of us has the ready capacity to look longer at another human and extract a more nuanced understanding of their motivations.

In his work, Humphrey uses the term consciousness to mean theory-of-mind-havingness. What I'm talking about is going into the world of narrative with an *awareness* of that consciousness, centered on the characters' minds (and how their theories change per unit of time) but extending in one direction to the author's mind and in the other to the mind of you, me, the reader. As a writer of fiction that tends to dwell in a moment and let the characters speak in

THEORY OF MIND & THE SCIENCE OF STORY

associative dialogue, theory of mind has become very important to me; it has massively enhanced my understanding of the subtleties of conflict, emotion, and insight.