Empathy, Narrative, and Art in Medical Education

Zachary Schwartz

Ethics and Medical Humanities Pathway
University of Miami Miller School of Medicine Class of 2020

Empathy training now holds a significant place in medical education. Many medical schools offer some form of empathy training as part of their communication skills curriculum, and examinees must express empathy to pass the clinical portion of Step 2 of the United States Medical Licensing Exam (Le 2018). Accordingly, there are hundreds of studies devoted to determining the best ways to teach empathy in medical students and residents (Sulzer et al. 2016). However, educators often fail to discuss what, precisely, empathy *is*, its nature and process. Indeed, the majority of studies on empathy training have "treated empathy itself as a black box, using global construct measurements that are unable to shed light on the underlying processes that produce empathic response" (Sulzer et al. 2016).

This omission is not surprising, because the question of what empathy *is* appears to be surprisingly difficult. While many people would likely claim an adequate first-hand understanding of empathy, its elaboration has proved surprisingly difficult for philosophers and psychologists (Coplan and Goldie 2011). That debate, which dives headlong into aesthetics, phenomenology, and philosophy of mind, has reached a level of specialization where it may be difficult for the average physician to participate. While this is unlikely to be an issue on an individual level—physicians do not need a stamp of approval from philosophers and psychologists to be empathic—neglecting the inner workings of empathy does seem to present a problem for any educational program that wishes to systematically transmit such capacities to students.

In this paper, I draw attention to the philosophical complexities of empathy, as they appear in medical education. At the most basic level, we encounter an overly functional approach to empathic understanding, where behaviors and rote speech are emphasized at the expense of any "inner" experience. Even where the goals of empathy training are more holistic, we find philosophical problems in some of the major scales used to assess empathy interventions. On the one hand, these scales seem to do a poor job operationalizing the concept of empathy; on the other hand, they imply certain beliefs about empathic processes (i.e. simulation theory) that have been heavily criticized. After detailing these criticisms, I summarize an alternative view of empathy based in narrativity. The narrative account helps to address the so-called "diversity problem" of empathy, i.e. how we are able to relate to those very unlike ourselves. It also accords nicely with the rise of narrative medicine and the push for literature and the arts in medical education. Nevertheless, a narrative theory of empathy still faces certain philosophical and practical difficulties, particularly in medical contexts, where the disparities between care providers and patients are often extreme. I attempt to address these problems through reference to literature and the arts, eventually defining a concept of "artful narrative." Though this concept may prove difficult to operationalize, it

is promising as a framework for understanding the role of narrative and the arts in medical education.

Functionalist approaches to empathy

Hesitancy to address the philosophical and psychological dimensions of empathy tends to encourage a superficial and highly functional approach to empathy. In this case, the *form* of empathy is emphasized without reference to its "inner" content, i.e. the feelings and mental processes of the empathizer. Such functionalism is apparent in the widely-taught SPIKES protocol for giving bad news to patients (Baile et al. 2000). The following excerpt is from step 5, "E—Addressing the Patient's Emotions with Empathic Responses":

> When patients get bad news their emotional reaction is often an expression of shock, isolation, and grief. In this situation the physician can offer support and solidarity to the patient by making an empathic response. An empathic response consists of four steps (Ptachek 1996):
> • First, observe for any emotion on the part of the patient. This may be tearfulness, a look of sadness, silence, or shock.
> • Second, identify the emotion experienced by the patient by naming it to oneself. If a patient appears sad but is silent, use open questions to query the patient as to what they are thinking or feeling.
> • Third, identify the reason for the emotion. This is usually connected to the bad news. However, if you are not sure, again, ask the patient.
> • Fourth, after you have given the patient a brief period of time to express his or her feelings, let the patient know that you have connected the emotion with the reason for the emotion by making a connecting statement.
>
> An example:
> Doctor : I'm sorry to say that the x-ray shows that the chemotherapy doesn't seem to be working [pause]. Unfortunately, the tumor has grown somewhat.
> Patient : I've been afraid of this! [Cries]
> Doctor : [Moves his chair closer, offers the patient a tissue, and pauses.] I know that this isn't what you wanted to hear. I wish the news were better.

The style of this passage—its awkward mechanical description of how we relate to others—offers a clue to the implicit philosophical view presented here. It begins with the cursory claim that the "empathic response involves four steps," as if this were an empirical fact, as the cardiac conduction cycle involves four steps or glycolysis ten. Of course this is not the case—the debate among philosophers and psychologists about empathic response is still very much alive, with no agreement whatsoever that it involves four discrete steps or even a series of steps at all. Indeed, the original paper (Ptacek 1996) recommends these four steps not as a model of what empathy *is* but simply as behaviors for delivering bad news (behaviors which are, by the authors' admission, "in need of empirical work"). In this way, the SPIKES protocol blurs the

distinction between empathy itself and those responses we associate with empathic feeling. Hence, the passage reads not as a description of what *actually occurs* in an empathic response, replete with the empathizer's own thoughts and feelings, but like instructions for the robotic simulation of empathy. All that is required to run the empathy "program" (assuming our possession of a certain facial recognition software) is: the input of a patient's emotion, a probabilistic guess at the reason for the emotion (which can be clarified by a few pre-programmed questions), the integration of the emotion and the reason for the emotion, and the output of a "connecting statement." The SPIKES protocol does not rely on any "inner" experience of empathy for the practitioner.

Such approaches to empathy training recall the famous "Chinese room" thought experiment by John Searle, in which we are asked to imagine a man, who does have any native grasp of Chinese, placed in a room and given a guidebook for responding to one Chinese symbol with another, according to empirical rules of standard conversation (Searle 1984). Written "input" statements are fed through one slot and the man, using his guidebook, feeds "output" statements back through another slot. From outside the room, it looks as though the man knows Chinese—but of course he does not. He has no genuine knowledge of the language, its meaning and context, but only of the functional rules for responding to one symbol with another (Kim 2011). Searle's argument is that this kind of functionalist "thinking," which is how basic computers operate, is not an appropriate model for human consciousness. The same can be said for those functionalist models of empathy like the SPIKES protocol. Just as the man in the Chinese room does not understand Chinese, so the rendition of certain statements and behaviors doesn't entail genuine empathy. While these behaviors may have a familiar function and form, and may provide some reassurance, they are clearly not sufficient for empathy itself. They are merely a procedure one undertakes, following a set of language rules, to elicit a desired response. By the same token, functionalist approaches to empathy ought to raise eyebrows in a field so eager to demonstrate its irreplaceability by artificial intelligence. While many physicians insist on their own value, over and against pure technology, because of certain "human qualities" inherent in expressions of empathy and care ("5 Reasons Why Artificial Intelligence Won't Replace Physicians" 2018), SPIKES and similar methods beg to differ. Indeed, in providing a series of rote steps to be followed by an anonymous practitioner, the SPIKES protocol may *itself* be considered a kind of technology (Sadler 2008).

On the other hand, we may be inclined toward a more charitable reading. Physicians confront countless cases of suffering every day, and to expect that they feel genuine empathy at every moment is unreasonable. Resources like the SPIKES protocol—which provide a repeatable formula for an empathy-like response, regardless of the physician's inner state—help ensure that patients' complaints are not met by a cold, blank stare. Such protocols also implicitly address a problem that will become thematic in the rest of this paper: what about cases when the physician *cannot* easily empathize, because the suffering demonstrated is beyond anything she could possibly identify with on a personal level? How can the average physician (who, by virtual necessity, has known a certain amount of success, wealth, social prestige, and so forth) empathize with the sheer variety and depth of suffering that comes through the healthcare system? I will return to this issue shortly.

Genuine empathy: simulation versus narrative

Of course there are other empathy curricula that make a more substantive claim to improve empathy, not just by providing a series of rote behavioristic steps, à la SPIKES, but to improve practitioners' capacity for genuine empathy (Sulzer 2016). Yet these curricula raise an interesting problem. In order to prove the effectiveness of an empathy intervention, there must be a way to measure empathy as an outcome. But if we take seriously the debate among philosophers and psychologists, we simply don't have a firm consensus on what empathy is. How, then, are these interventions assessed?

There are many scales that claim to assess empathy, of varying quality (Hemmerdinger 2007). Rather than attempting a broad overview, I will focus my analysis on one of the most used and frequently validated scales, the Jefferson Scale of Empathy (JSE). The JSE, which has been given in 74 countries and translated into 56 languages, was developed specifically for healthcare providers and trainees, and thus accounts for much of we claim to know about empathy training both in the United States and abroad (as partial representation: De Lillo et al. 2009; Glaser et al. 2007; Hojat et al. 2013). While certainly not representative of all scales, the JSE captures some of the most common approaches to empathy in medical education.

The first problem is that scales like the JSE is that they tend to not operationalize empathy *itself* so much as whether the participant *values* (or claims to value) empathy. Questions assess whether the participant views empathy and empathic behaviors as "important," "helpful," whether they believe emotional understanding contributes to better outcomes and physicians "should try" to be more empathic (Hojat et al. 2001). The scale does *not* assess the extent to which participants actually exercise or are capable of such feelings and behaviors. Though not all scales so clearly demonstrate the split between the exercise of empathy and its valuation (for example, the Toronto Empathy Questionnaire), the issue may naturally arise through self-reporting/social desirability bias (King 2000).

It is not clear how closely tied, or in what way, these two states—being empathic and valuing empathy—really are. When Sophocles' Odysseus pities the broken Ajax, his enemy, the experience is clearly unwelcome, perhaps even shameful (1957). Nietzsche, perhaps the strongest critic of the *value* of compassion and sympathy in the Western Canon, was reportedly quite soft-hearted himself, often worrying whether he had offended acquaintances and, in the famous episode in Turin, at the precipitation of his "madness," breaking down at the sight of a draft horse being whipped (Cybulska 2015). The converse seems even more common: we often encounter people who adamantly express their belief in the value of empathy but are—as evidenced by subtle observation—narcissistic, self-protective, manipulative, and even cruel. In fact, it often appears that such people (who would no doubt score highly on the JSE but in a parallel universe might turn informer for the gestapo) are unknown to themselves, genuinely believing in their own superior empathy and virtue. This observation, more sophisticated and generalized, has given rise to some of the more biting critiques of compassion and sympathy in the Western canon (Nietzsche 1997). Indeed, it may come as a surprise to

medical educators that there is a strong philosophical tradition in the West of questioning the overvaluation of compassion and sympathy (Nussbaum 1994).[1]

Even if we do take for granted the link between the conscious valuation of empathy and empathy itself, scales like the JSE still raise significant philosophical problems. While the JSE never explicitly defines empathy, it does imply a certain set of beliefs about the way empathy works. These beliefs are expressed in the following three items from the JSE:[2]

> Item 17: Physicians should try to think like their patients in order to render better care.
> Item 9: Physicians should try to stand in their patients' shoes when providing care.
> Item 15: Empathy is a therapeutic skill without which the physician's success is limited.

For the sake of further analysis, let us abstract from the medical context and restate these beliefs in a more general form, as an attitude toward empathy itself: First, empathy is taken to be deliberate and intentional. One must "try"; it takes *work*. Second, the process by which empathy occurs is *simulation*, in which we try to envision the thoughts and feelings of another as if they were our own. Third, empathy is a *skill* employed by the empathizing individual, an instrumental process whose success depends largely on the efforts and talents of that individual.

These interrelated beliefs about empathy, embodied in the JSE, cohere into what is known by philosophers and psychologists as "simulation theory." This theory—that empathy occurs through the process of simulating the thoughts and feelings of another—is perhaps the default position of the general public regarding empathy, though it has academic support as well. In fact, there are many versions of simulation theory, with some proposing that empathy is no different than ordinary social understanding (low-level theory) and others that empathy is a special case of social understanding, requiring a deliberate effort to simulate the thoughts and feelings of another (Goldman 2006). The picture implied by the JSE—with its emphasis on a deliberate and conscious effort to take the patients' perspective—is more in line with high-level simulation theory. This also reflects the ethos of the medical profession, which tends to emphasize the value of hard work, deontological ethics ("one ought try to be empathic"), and technical mastery.

Yet despite the intuitive appeal of this theory, and its academic support, it has suffered from a number of criticisms. Some of the most significant have come in the last

---

[1] We should note here that "empathy" is a relatively modern term. Though its etymology suggests roots in ancient Greece, it was actually brought into English in 1909 by the psychologist Edward Titchener, as a translation of the German "Einfühlung" ("feeling into"), which was important for aesthetics. Though there is much philosophical work trying to separating empathy from related concepts like sympathy, compassion, and pity, these words are rarely used with such precision in medical education. I have retained the same loose usage in this paper. While empathy, narrowly understood, is still my main concern, there is such significant and unresolved overlap between these concepts that I have not devoted significant space to parsing out the precise differences between them.

[2] These are analogous to items on the Interpersonal Reactivity Index (IRI). The IRI is an older empathy scale, developed for the general population, whereas the JSE was specifically developed for healthcare providers (Hojat 2005).

decade or so, from a group of phenomenologists including Shaun Gallagher. Gallagher summarizes the main philosophical problems with simulation theory in a paper titled "Empathy, Simulation, and Narrative" (2012). As these are particularly relevant to the implicit role of simulation theory in medical education, I will here detail two of the main branches of this criticism (with some of my own additions to the argument).[3]

The first branch pertains to the supposed neurobiological evidence for simulation theory, i.e. "mirror neurons." To clarify, mirror neurons refer to those neurons that fire not only when we perform an action but when we witness the same action performed by another. This neurobiological phenomenon is often taken as evidence (particularly by physicians, who tend to prefer "hard science") that empathy is a process in which we *simulate* the mental processes of another. Yet this is a superficial understanding, which breaks down under close scrutiny. While mirror neurons may in fact be necessary for empathy, they are clearly not sufficient to describe the process. First, the phenomenon of mirror neurons is not as simple as most people imagine. Only about one-third of mirror neurons fire at a given time, and the phenomena is often one of complementary action more than direct matching (Csibra 2005; Newman-Norlund 2007). More philosophically, there is a disjunction between the kind of volitional, teleological language of simulation theory and the mechanics of neuroscience. Mirror neurons work at the behest of concentration gradients, action potentials, and other physiological processes—not conscious decision. Their firing is not initiated by deliberate choice, by the moral effort of "trying to put oneself in the patients' shoes," but by unconscious physiological processes. They were, after all, first discovered not in human being but in macaque monkeys, who ostensibly do not feel a moral compulsion to empathize.

To this I would add another problem, though Gallagher doesn't mention it: the conundrum of how neural activity translates into mental experience. This is the so-called "hard problem of consciousness," and it is also a problem for a simulation theory of empathy based on mirror neurons. Even if we assume a perfect correspondence between their brain scans, we have no way to verify that the two participants are sharing a genuinely similar mental state, because we do not understand how physiological phenomena become mental *experience*. That these brain scans are *not* precisely the same, reflecting (relatively weak) correspondence rather than direct matching, only strengthens this point. Furthermore, It's important to remember that the vast majority of mirror neuron experiments in human beings are carried out by fMRI, which merely illustrates blood flow to (and therefore increased activity of) a particular region. In other words, we are not seeing the resonance of individual neural circuits but of general areas. This is rather like observing, from satellite imagery, that the downtown districts of Miami, Florida and Lander, Wyoming both fill up on a Friday night, without knowing much about what people are doing there. Only a handful of studies have worked to get around this problem, and the results are mixed (Kilner 2013).

This leads us directly into a discussion of the next argument against simulation theory, the so-called diversity problem, which will become a theme for the remainder of the paper. The diversity problem concerns the sheer variety of personalities, situations, and modes of experience, and the difficulty of forming a genuine empathic connection

[3] Gallagher's other criticism, which I do not mention, concerns the problems of understanding empathy from a childhood-developmental perspective. While sound as a criticism of simulation theory, this is not as relevant to the present topic.

between these. How do we know, when we try to imagine the experience of a patient, that we are not merely projecting our own perspective onto theirs? Do we feel what they feel, or only a shallow, misguided representation?[4] This is, as it were, the common sense formulation of the diversity problem. It also has a logical corollary, which Gallagher calls the "starting problem." Simulation theory posits that, when we empathize, we imagine a mental state intended to match a target's mental state. But how do we know *what* will match the target, if we do not already understand their mental state? Whichever argument seems more convincing, one must concede there is here a serious problem with simulation theory. That problem concerns the difficulty of how we could ever simulate the thoughts and feelings of those who are truly different from us. Interpersonal similarity is, as Pierre Jacob suggests, "arguably the major assumption of a simulation-based approach to empathy" (Jacob 2011). What to make of empathy, then, in the context where doctors—whose young adult lives are generally defined by stability, social respect, academic success, and so forth—often have little in common with many of their patients?

Gallagher and other phenomenologists have tried to answer the diversity problem by positing an alternate view of empathic processes, where empathy depends not on willful simulation but on narrative competence. We are not isolated atomistic beings, phenomenologists argue, who must work to recreate the private mental space of others by sole reference to our own private mental state. Rather, we come to know others and ourselves through our mutual habitation in a shared world, and the telling and re-telling of stories of that world. These stories determine not just how we view the trials, setbacks, and motivations of others but how we developmentally come to understand ourselves. In that sense, they elaborate a shared world of meanings, a horizon that contains all. When we go through some event (a birth, a death, a marriage, a divorce), it is always already structured—in its very lived experience, as *we* experience it—by the narratives that pervade our shared existence. Thus we are not the sole owners of our experience, and understanding another human's world does not require simulating some other private and self-contained mind. As Karl Jaspers says, "We understand people not through considering and analyzing their mental life, but by living with them in the context of events, actions and personal destinies" (1912/1968). We can legitimately understand each other because, in short, we live in the same world, a world whose complexity and differences are given through narrative. Empathy is not the act of deliberately displacing ourselves into the headspace of another, mirroring their precise thoughts and feelings. Rather, we take part in a unique experience (a distinct "intentionality," in the language of phenomenology) in which we, *as ourselves*, relate to someone as a character in their own story.

This correlates nicely with recent attention to narrativity in the medical humanities. From the beginning, advocates of narrative medicine have intuited the connection between narrative competence and clinical empathy. Narrative knowledge "provides a rich, resonant comprehension of a singular person's situation as it unfolds in time, whether in such texts as novels, newspaper stories, movies, and scripture or in

[4] Nietzsche gives an early version of this diversity argument in *Daybreak*, playing on the etymology of *Mitleid*, the German parallel of the English *compassion* or *sympathy* (in all cases, German/Greek/Latin, the roots are "with" + "suffering"). He says: "It is misleading to call the *Leid* (suffering) we may experience at such a sight, and which can be of very varying kinds, *Mit-leid* (pity) for it is under all circumstances a suffering which he who is suffering in our presence is free of: it is our own, as the suffering he feels is his own" (Nietzsche 1997).

such life settings as courtrooms, battlefields, marriages, and illnesses," as Rita Charon says, arguing for the inclusion of narrative training in medical education (2001). However, while this connection—between a more philosophically sound model of empathy on the one hand, and the resources of the medical humanities on the other—is very appealing, I would like to consider some potential objections to a narrative account of empathy in a medical context. Though I ultimately support a narrative account—and alongside it, the importance of the arts in cultivating empathy among healthcare professionals—clarification of the relevant features is necessary.

Objections to a narrative account of empathy

There are lingering philosophical questions about Gallagher's narrative account of empathy. For one, it's not clear that this account has escaped the diversity problem. How do we know that there aren't other narratives beyond our understanding? How would we recognize when we saw one, rather than just reframing it in terms of our own narrative? Certainly there are some modes of suffering (as well as joy, which can also be an object of empathy, though I've underemphasized it here) that are so extreme as to not be adequately expressed by the dominant cultural narratives. This is an argument often expressed on behalf of marginalized social groups, who must guard against hegemonic interpretations of their identity (Somers 1994). But the point applies to individual examples as well and can be found on virtually any floor of the hospital. Can we even begin to understand the narrative of (and therefore properly empathize with) a mother who's lost her child, if we've never lost a child ourselves, if we've never even had children? Can we genuinely understand narratives of addiction if we've never felt its pull?

Gallagher tries to head off this question by appealing to the importance of "communication skills" in revealing a narrative. When we do not understand we ask questions, we listen, we follow up, we clarify. He does not say much more than this, only briefly addressing the issue in a footnote as a response to a reviewer (Gallagher 2012, 371). Neither does his response address, of course, the particular issue of doctor-patient relationship. At the very least, it's clear that communication skills are not a *procedure*, deliberately undertaken by a physician, to "get inside the head" of a patient (as simulation theory would have it), but a naturally unfolding conversation between human beings. Empathy is not a "therapeutic skill" of a single provider, as the JSE would have it, but arises from the dynamic interaction between two individuals.[5] What

[5] This dynamic interaction is often lost in formal "empathy exams," such as the Communication Skills component of Step 2 of the United States Medical Licensing Exam. Here, standardized patients are instructed to "stick to the script" as examinees ply them with empathic statements, trying to elicit information necessary for the diagnosis. Examinees also receive points, necessary for a passing score, each time they verbally "express empathy" in the course of the 15-minute simulation (Jamison 2014). Standardized patients are sometimes instructed to remain intransigent and difficult, to press the students to demonstrate empathetic resolve (Le 2018). In essence, "communication skills" come to be treated very literally, as a series of techniques mastered by a single individual which can be exercised to help achieve a definite end. Empathy is viewed as a *literal procedure*, like bronchoscopy or laparotomy, in which one learns the techniques to "open up" difficult patients to further questioning (or perhaps achieve some other end, such as pacification or reassurance). This approach—procedural, repeatable, unilateral—is a far cry from the kind of communication skills Gallagher describes as necessary for narrative and empathy. Notably, what is required for the latter is *time*. Unlike the procedural approach to empathy—whose display, during a 15-minute patient simulation, seems to depends solely on the skill of the provider—an organic, naturally unfolding

this does seem to require, however, is a certain narrative capacity of those individuals; in order for such conversations to take place, they must have not only the ears for it but the voice as well. This problem will become significant in the case of severe mental illness, discussed below.

Before tackling that problem, however, it's worth addressing a more general concern about narrativity and empathy. This is the question of whether narrative thinking is really so fundamental to human life, and may not *itself* limit our understanding of others. That is, narrativity may just be one among many modes of self-understanding, and in that sense may simply be another way in which we egocentrically project our own life onto others. Galen Strawson writes, in "Against Narrativity":

> There are deeply non-Narrative people and there are good ways to live that are deeply non-Narrative. [The beliefs that narrativity is universal and/or crucial to the good life] hinder human self-understanding, close down important avenues of thought, impoverish our grasp of ethical possibilities, needlessly and wrongly distress those who do not fit their model, and are potentially destructive in psychotherapeutic contexts. (2004)

Strawson's argument in "against narrativity" is largely based on his own personal experience of non-narrativity, as well as the first-person reports of fellow non-narratives. Angela Woods extends this criticism specifically to narrative medicine, arguing that quest narratives of the "triumphant journey through illness" can obscure the self rather than reveal it. (Woods 2013). While I do not fully agree with Woods' wholesale criticism of narrative (for reasons I will detail below), I think there is something salient in it. Many illness narratives *are* marked by the repetition of certain narrative themes, which can represent a rather simplistic view of human experience: heroism, linear growth, communion with others, the triumph of good over evil; these as opposed to the chaos of the "broken, interrupted, distressing and ultimately ineffectual" (Woods 2013, 122). While we should never discourage patients or physicians from trying to formulate their struggles into a story, we ought to remain aware of the storyteller's own literary or narrative capacity, and how this can promote or inhibit empathy. To tell a *good* story, one that inspires deep understanding, often means breaking with the clichés and tropes that mark many narratives as predictable and forgettable. Such formal qualities, I would propose, are often more important to the impact of a narrative than the bare events that compose it.

This same issue surfaces in a slightly different way when we consider the unique relationship of narrative and illness. Matthew Ratcliffe, another phenomenologist, has written extensively about the particular difficulties of empathy with psychiatric illnesses. Psychiatric patients, he notes, frequently report failures of empathy in those around them (Ratcliffe 2017, 2). While it's difficult to clarify the reasons for this (since those with certain disorders are also prone to interpret positive or neutral gestures as hostile), it seems clear that mental illness represents not just an intensifying of everyday emotions but the loss of a shared background world. This shared world, now lost, is what allowed

---

conversation cannot be rushed. Perhaps this is why there appears to be such crisis of empathy in medicine, at the moment when the average doctor's visit is also around 15 minutes (Tai-Seale 2007).

for genuine mutual understanding in Gallagher's theory, over and against attempts at simulation. Hence psychiatric disorders seem to present a problem for our narrative theory of empathy, at least as previously described. Those without mental illness simply cannot comprehend the "isolated, alien realm that is set apart from the consensus reality taken for granted by others as an unwavering backdrop to their experiences, thoughts and activities" (Ratcliffe 2017, 5). The tendency of narrative medicine to overlook psychiatric disorders seems to support this phenomenological gap (Woods 2013, 120).

In response to this problem, Ratcliffe makes a crucial point: genuine empathy often requires not just the understanding of interpersonal similarity (again, the major assumption of simulation theory) but the recognition of *radical difference*. In fact, "when acknowledgement of difference by A is successfully conveyed to B, this sometimes suffices for B's recognition of empathy on the part of A" (Ratcliffe 2017, 1). We should note this mode of empathizing is relatively common in medicine and extends to situations beyond psychiatric illness. Phrases like, "I can't even imagine," or "I have no words," are familiar to most physicians and serve to signify the recognition of profound phenomenological difference. They represent an appropriate distance from which the physician—regardless of her own background—can be assured she will not overstep her understanding, misinterpreting or minimizing the suffering of her patient.[6] To be clear, Ratcliffe still endorses the importance of conversation and narrative in clarifying these radical phenomenological differences. In a convincing analogy, he compares illness narratives with the experience of looking at a piece of art while it's described by someone with knowledge of art history. Indeed, by listening to the story of how the artwork was made, its influences, its historical context, and so on, the piece does come to *look* differently. The same process holds, in Ratcliffe's view, in dialogue with patients, who are the historians of their own sufferings (Ratcliffe 2017, 12).

Still, one wonders how a narrative approach to empathy holds up at the extremes of psychiatric or neurologic disorder. Here, with patients who are by definition "poor historians," conversation and narrative can be fractured or obliterated. How is it possible to grasp the suffering of a psychotic patient when the narrative is tangential, disorganized, or utterly unrecognizable? How can one empathize with depression so severe that it blunts speech or drifts into catatonia? While these are extreme examples, they point to an important issue that already arose peripherally with Strawson's "anti-narrativity" claim: What of the *capacity* to tell an accurate, original, and impactful story? I will address this idea in the next section, where I consider the often discussed but poorly understood relationship between empathy, narrative, and art.

Narrative capacity and artful narrative

For clarity, let me briefly summarize the argument so far. Empathy training has become an increasingly important part of medical education, but, despite an abundance of research on the best means for teaching empathy, there is little discussion of its philosophical intricacies. This omission tends to result in an overly functional approach

[6] Unfortunately, these phrases can also become an absentminded reflexive habit, a barrier to empathy. Statements like, "that must be very hard for you," when said without genuine attention to the suffering at hand, can simply reflect disregard, the desire to escape the moment.

to teaching empathy. Functionalist training paradigms, exemplified by SPIKES protocol, represent at best a stunted view of empathy and at worst an imitation of empathy, the rote learning of empathy-like behaviors. While some of these empathy-like behaviors are necessary for physicians, they shouldn't be confused with empathy itself. Even curricula that do claim to teach genuine empathy should be carefully scrutinized, given the scales on which they must be validated. On the one hand, these scales do not seem to measure empathy itself but the *valuation* of empathy; on the other hand, they can imply a misguided picture of empathic processes. In the case of the often-used JSE, this picture is the so-called simulation theory of empathy, which has received heavy criticism from many philosophers. The most significant of these criticisms, especially in the context of clinical medicine, is the diversity argument, which suggests the difficulty of simulating the experience of another person, especially a person who is significantly different from ourselves. This has particular relevance for the doctor-patient relationship, as most doctors are, almost by definition, competent, successful, intelligent, and emotionally well-regulated, while many of their patients are not. In place of the simulation theory, philosophers like Gallagher and Ratcliffe have promoted the importance of narrative for empathy. This seems in many ways an improvement over simulation theory, and allows for the paradoxical insight that genuine empathy can require the recognition of profound difference. A narrative theory of empathy also correlates nicely with the recent focus on narrative in the medical humanities. Nevertheless, a narrative theory of empathy raises some difficult questions. These concern the ways narrative empathy can break down, either because of 1) people whose lives are supposedly not well-captured by narrative, and can even be misunderstood by narrative or 2) when the capacities for narrative are lost, as in the case of certain psychiatric disorders. I will address these concerns below. My proposal is that it is precisely regarding these problems that the humanities (I have focused on literature here, though other forms are open to discussion) show their chief value for medical education. However, the very manner in which they are valuable makes them difficult to study by quantitative empirical methods.

To begin, let us approach the concept of narrative itself. Like empathy, narrative is everywhere in human life yet difficult to precisely define. Though most people would claim to understand narrative, it is subject to significant academic debates. In common understanding, narrative is linear, diachronic, coherent: a unified character moving through a beginning, middle, and end (Keen 2006). This is the kind of narrativity that Strawson rejects as being a universal feature of human existence, and which he cites as a potential limitation to our understanding of others. Yet this seems to be an overly restrictive definition, depicting one kind of narrativity but not narrativity itself. As such, we must understand Strawson's criticism as focused on a particularly narrow mode of storytelling. Strawson would likely agree that people have stories (not of the "Grand Life Story" sort, but simply "stories") that are their own, and that in telling these stories people come to reveal something important about their experiences, something that often marks their experiences as different in radical ways. Indeed, this is the very basis of Strawson's own argument: his recounting of his own life, that it does *not* cohere in simple diachronic fashion, and that his own impressions do not square with what he understands *from other peoples' narratives*. In essence, it is only on the basis of narrativity, broadly defined, that Strawson is able to make his argument "against narrativity." It's also telling that Strawson cites a number of narrators who *do* capture the

kind of “episodic” (vs. “diachronic”) experience he’s referencing: Updike, Woolf, Conrad, Borges, Pessoa (who, it’s relevant to note, wrote under more than 75 separate heteronym identities), and many more. Though authors are clearly atypical in their approach to narrative, it would be misleading to call them non-narrative. We might push the point even further in reference to the other arts. Is it appropriate to exclude the visual arts (which lack explicit temporal structure) or music (which lacks character or point-of-view) from conceptions of narrative? In any case, the boundaries of narrative are wide and indeterminate.

This broad conception appears to “save” narrativity from the Strawsonian critique, but it also reveals why the relationship between narrative and empathy is difficult to quantify by empirical research. Previous attempts to connect simple narrative structures—first person point-of-view, emotionally open characters, etc.—with predictable empathic responses have generally proved unsuccessful or overly complex (Keen 2006). One wonders, then, how it would ever be possible to capture and quantify the effects of such stylistically diverse, convention-bending authors as listed above, authors who often inspire deep feelings of attachment and are capable of bringing tears to the eyes of their readers. And what of those backgrounded narrative features which are hard to delineate, inseparable from the whole, such as mood or tone? These “poetic” features are no doubt hugely important to a proper understanding of narrative and empathy, yet they are difficult to operationalize. As the critic Orville Prescott said of Truman Capote’s famously eerie first novel, shrouded in themes of illness and decadence, “It is not possible to be certain just what is going on in certain passages of ‘Other Voices, Other Rooms’…But is impossible not to succumb to the magic of his writing…as still and ominous and flickering with unholy light as the last five minutes before a summer cloudburst” (1948). How does one operationalize this? Such “succumbing” is, as Prescott notes, quite “magical,” for it seems relatively unrelated to what’s “going on” in terms of more concrete events of the narrative. Yet this property—which we might call the *style* of a story, the very *art* of a piece of art—is unquestionably bound up in the features of narrative we care about in regard to empathy. Is it not this very artistry, above and beyond the basic narrative components present in *all* stories, that makes us “succumb” to a particular text, illuminating a world that seems radically different than our own?

Hence this issue turns out to be intimately connected to the second problem, of empathy and otherness. As stated above, the problem of narrative *capacity* is revealed most clearly in psychiatric or neurologic illness, where people not only inhabit a radically different world but often lack the means to articulate it. Yet there are important exceptions to this situation, and these are often revealed most clearly in the arts. Here one can reference Thomas Mann, obsessed as he was by the competing destructive and creative energies of disease:

> Disease…first of all it is a question of who is sick, who is insane, who is epileptic or paralytic: an average dolt, whose disease, of course, lacks all intellectual and cultural aspects—or a Nietzsche, a Dostoevsky. In their cases the disease bears fruits that are more important and more beneficial to life and its development than any medically approved normality. (1945)

If one can look past the elitism, Mann's observation speaks well to the relationship of narrative, empathy, and otherness. He notes the ability of disease, when channeled through "the artist," to open up new doors of thought, to lay out unseen narratives; such artists may even—though we must be cautious with such claims—speak for those who lack the capacity. While we must be wary of romanticizing this phenomenon, there is no question that artists of a particular caliber are able to use disease to their advantage, revealing a world that often remains hidden from view.[7] Here is Virginia Woolf—a sick soul if there ever was one, presumably bipolar and occasionally psychotic—describing this phenomenon in her fabulous essay *On Being Ill*:

> How common illness is, how tremendous the spiritual change that it brings, how astonishing, when the lights of health go down, the undiscovered countries that are then disclosed, what wastes and deserts of the soul a slight attack of influenza brings to light, what precipices and lawns sprinkled with bright flowers a little rise of temperature reveals, what ancient and obdurate oaks are uprooted in us in the act of sickness, how we go down into the pit of death and feel the waters of annihilation close above our heads… (1926/2002)

While illness may be common, as Woolf says, those who can write about with her skill are not. Psychiatric and neurologic illness reveals that the transparency between how one feels and what one can express is not a given, and the "sick artist" further reveals that this not an either/or phenomenon. Some narrators are simply *better* at articulating their lived experience of illness. Therefore, our understanding of that experience depends not only on us but on them as well, on the ill person's skill at bringing their suffering into view. Again, this requires appreciating, contra the kind of simulation theory implied by the JSE, that empathy is not a willful act on the part of the empathizer. Even those patients without severe neurological or psychiatric deficits may fail to elicit our understanding, if they do not have a certain narrative capacity to make themselves heard. The superior development of such narrative capacities is, no doubt, what separates someone who merely knows how to communicate from being an *artist*. "Most people are affected by an inability to say what they see or think," as Pessoa says in *The Book of Disquiet*. "Impressions are incommunicable unless we make them literary" (2001).

This is, I put forward, is among the greatest potential benefits of art in medical education. But the crucial point is that this must be art of a particular kind: those artworks of a particular quality that are able to "show, not tell" of the kinds of phenomenological difference that characterize experiences of illness and suffering. This art ought to be considered broadly "narrative" in the sense I've laid out above, in that it makes genuine empathy possible by communicating a certain world, but is perhaps defined more by the way it violates conventional narrative forms than by how it follows them. Let us call this *artful narrative*.

---

[7] I do not mean to limit this to established or canonical artists. There are clearly testimonies from psychotic and depressed patients that strike us with a great deal of "art" and to which we, for a variety of reasons, empathically "succumb." Yet these are not published texts, generally speaking, and they are therefore difficult to reproduce for study. The point is to highlight the capacity for description and articulation, which can occur anywhere but is naturally a central feature of good art and literature.

Unfortunately, one problem with this notion of artful narrative is that even if we could agree precisely what it was and what features it entailed, it seems exceedingly difficult to operationalize for controlled study. Yet we cannot deny that certain narrators have a preternatural capacity, above and beyond a basic retelling of events, to give voice to new and different worlds, to describe worlds that are no doubt well populated but cannot be easily articulated. Such narrative capacities must be of immense interest to anyone who would study empathy in those "undiscovered countries" of illness.

Conclusion

To recount, empathy is a difficult and nuanced concept, and many of its philosophical problems show up in the complicated and sometimes contradictory attempts to teach medical professionals to be more empathic. One of the greatest challenges in understanding (and perhaps teaching) empathy is the difficulty of grasping, in a deep sense, lives very different than our own. At the academic level, this problem is captured by the diversity argument against simulation theory. As an alternative to simulation theory, phenomenologists like Ratcliffe and Gallagher have proposed the importance of narrativity. I examined two potential challenges to a narrative view of empathy, one based on the "anti-narrativity" of Strawson and the other based on the ways that certain psychiatric or neurologic illnesses can limit narrative capacity. This led not the wholesale rejection of narrativity but to an elaboration of what I called artful narrative. Well-written, "artful" literature (and perhaps other arts) may prove useful in articulating the kinds of radical phenomenological differences characteristic of great illness and suffering.

But I want to be cautious with this claim. One must not be too eager to project particular narratives onto particular patients. It would be a mistake to automatically interpret the man with epilepsy in bed 214 by the words of Dostoevsky, or the suicidal woman in bed 308 by the words of Woolf. The point is rather that literature and the arts can crack open the very possibility of difference, demonstrating what it's like to sink into a narrative that's radically different in structure as well as content. My argument is not that, by mastering this or that text, one becomes a master of this or that empathic situation, and so by mastering more texts one could gradually come to master empathy itself. Such thinking is far too procedural and systematic. Both functionalist and simulation-based curricula treat empathy as technical skill, to be produced by repeatable formula in class after class of trainees. This reflects a greater overall tendency in medicine, one to which even the profession's exemplars of "humanism" are prone (see Atul Gawande, 2012, on making medicine more like the Cheesecake Factory). I mean to resist this approach, at least with respect to empathy training. As such, I've said relatively little about practical details of which art we should promote, or how often, or how we verify its effects. I simply do not know these details, nor do I know how one might establish them in an "evidence-based" manner. There are, of course, many studies which attempt to prove the importance of literature and the arts to empathy. Yet these cannot avoid similar problems in operationalization as functionalist and simulation-based accounts (even if the underlying theory is a philosophical improvement). If we add to this the difficulty of grasping those particular "artistic" features that lead to us "succumb" to one story rather than another, the ability of empirical scientific studies to explain the true nature of narrative empathy is doubtful.

However, the connection between empathy and the arts ought still be of profound interest to medical educators. It may be that artful narrative offers the best chance for cultivating empathy in medical trainees, whose own lives and background worlds often differ markedly from their patients'. Having a framework to conceptualize this process is important, even if it doesn't easily conform to empirical scientific study.

What I'm offering in this paper, then, is mostly critique: of the overly technical treatment of empathy in medical training; of the notion that empathy (in accord with simulation theory) is just another opportunity for self-effacement and hard work on the part of physicians, rather than a natural and dynamic arising between individuals; of the tendency to treat art and narrative in simple, mechanical, procedural terms. I do not mean to construct a new training program so much as bring to light the philosophical difficulties of what, to many in medical education, must seem a skill to be learned like any other. Still, I do not want to be too harsh in this criticism. The fact that empathy and communication skills have become a focus of medical education is clearly an advance over the exclusive focus on technical and scientific knowledge. We just need to be careful to not unreflectively treat the former under the guise of the latter.

Finally, we must keep in mind that empathy is fundamentally a relation between living beings, not between humans and books, nor between humans and computer simulations. It cannot be produced, ex nihilo, out of art (even great art) any more than out of the SPIKES protocol, and even the best works are only a supplement to our face-to-face encounter with those beings like (or unlike) us. There is always something mysterious and ineffable about these encounters, which is what makes empathy a philosophical problem in addition to a practical one. The issue is not just that some physicians lack empathy, or that we do not know the proper strategies for making physicians more empathetic. In the final analysis, despite our own personal experiences of empathy, we do not share an explicit consensus on just what it is or how it works. Even the narrative theory of empathy is subject to significant internal debate. And so, as long as medical education takes upon itself the goal of systematically training doctors to be more empathic, philosophical questions surrounding empathy will keep rising to the surface.

References

Baile, Walter F., Robert Buckman, Renato Lenzi, Gary Glober, Estela A. Beale, Andrej P. Kudelka. 2000. "SPIKES—A Six-Step Protocol for Delivering Bad News: Application to the Patient with Cancer." *Oncologist*. 5(4): 302-311.

Charon, Rita. 2001. "Narrative Medicine: A Model for Empathy, Reflection, Profession, and Trust." *JAMA*. 286(15): 1897-1902.

Coplan, Amy and Peter Goldie. 2011. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.

Csibra, Gergely. 2005. Mirror neurons and action observation: Is simulation involved? In: *What do mirror neurons mean?* Available at: https://pdfs.semanticscholar.org/f1ec/4f02190061b90ca531d900c22e413bc66015.pdf

Cybulska, Eva. 2015. "Nietzsche's Übermensch: A Glance behind the Mask of Hardness." *Indo-Pacific Journal of Phenomenology*. 15(1): 1-13.

De Lillo, Mariangela, Americo Chiccheti, Alessandra Scalzo; Francesco Taroni, Mohammedreza Hojat. 2009. “The Jefferson Scale of Physician Empathy: Preliminary Psychometrics and Group Comparisons in Italian Physicians.” *Academic Medicine*. 84(9): 1198-1202.

“Five Reasons Why Artificial Intelligence Won’t Replace Physicians.” 2018. *The Medical Futurist*. https://medicalfuturist.com/5-reasons-artificial-intelligence-wont-replace-physicians

Gallagher, Shaun. 2012. “Empathy, Simulation, and Narrative.” *Science in Context*. 25(3): 355-381.

Gawande, Atul. 2012. “Big Med.” *New Yorker*. Available online at: https://www.newyorker.com/magazine/2012/08/13/big-med

Glaser, Karen M., Fred W. Markham, Herbert M. Adler, Patrick R. McManus, Mohammadreza Hojat. 2007. “Relationships between scores on the Jefferson Scale of physician empathy, patient perceptions of physician empathy, and humanistic approaches to patient care: A validity study.” *Medical Science Monitor*. 13(7): CR291-294.

Goldman, Alvin. 2006. *Simulating minds: The philosophy, psychology and neuroscience of mindreading.* Oxford: Oxford University Press.

Hemmerdinger, Joanne, Samuel Stoddart, Richard Lilford. 2007. “A systematic review of tests of empathy in medicine.” *BMC Medical Education*. 7:24.

Hojat, Mohammedreza, Salvatore Mangione, Thomas J. Nasca, Mitchell J.M. Cohen, Joseph S. Gonnella, James B. Erdmann, Jon Veloski, Mike Magee. 2001. “The Jefferson Scale of Physician Empathy: Development and Preliminary Psychometric Data.” *Educational and Psychological Measurement*. 61(2): 349-365.

Hojat, Mohammedreza, Salvatore Mangione, Gregory C. Kane, Joseph S. Gonnella. 2005. “Relationships between scores of the Jefferson Scale of Physician Empathy (JSPE) and the Interpersonal Reactivity Index (IRI.” *Medical Teacher*. 27(7): 625-628.

Hojat, Mohammedreza, David Axelrod, John Spandorfer, Salvatore Mangione. 2013. “Enhancing and sustaining empathy in medical students.” *Medical Teacher*. 35(12): 996-1001.

Jacob, Pierre. 2011. “The Direct-Perception Model of Empathy: a Critique.” *Review of Philosophy and Psychology*. 2: 519-540.

Jamison, Leslie. 2014. *The Empathy Exams*. Minneapolis. Graywolf. 1-26.

Jaspers, Karl. 1912/1968. “The Phenomenological Approach in Psychopathology.” *British Journal of Psychiatry*. 114: 1315.

Keen, Suzanne. 2006. “A Theory of Narrative Empathy.” *Narrative*. 14(3): 207-236.

Kilner, J.M. and R.N. Lemon. 2013. “What We Know Currently about Mirror Neurons.” *Current Biology*. 23(23): R1057-R1062.

Kim, Jaegwon. 2011. *Philosophy of Mind*, 3rd Ed. Boulder: Westview Press. 164-165.

King, Maryon and Gordon Bruner. 2000. “Social desirability bias: a neglected aspect of validity testing.” *Psychology and Marketing*. 17(2): 79-103.

Le, Tao, Vikas Bhushan, Kachiu C. Lee, Maniver Deol. 2018. *First Aid for the USMLE Step 2 CS*. Sixth Edition. New York: McGraw Hill. 25-26.

Mann, Thomas. 1945. “Dostoevsky—in Moderation.” In *the Short Novels of Dostoevsky*. New York: Dial Press. vii-xx.

Newman-Norlund, Roger D., Matthijs L. Noordzij, Rudd G.J. Meulenbroek, and Harold Bekkering. 2007. “Exploring the brain basis of joint attention: Co-ordination of actions, goals and intentions.” *Social Neuroscience*. 2(1):48–65.
Nietzsche, Friedrich. *Daybreak: Thoughts on the prejudices of morality*. Edited by Maudemarie Clark Clark and Brian Leiter. Translated by R.J. Hollingdale. Cambridge: Cambridge University Press. 83-85.
Nussbaum, Martha C. 1994. “Pity and Mercy: Nietzsche’s Stoicism.” In *Nietzsche, Genealogy, Morality*. Berkeley: University of California Press. 139-167.
Pessoa, Fernando. 2001. *The Book of Disquiet*. Edited and Translated by Richard Zenith. New York: Penguin. 107-108.
Prescott, Orville. 1948. http://movies2.nytimes.com/books/97/12/28/home/capote-voices.html
Ptacek, J.T. and Tara L. Eberhardt. 1996. “Breaking bad news: a review of the literature.” *JAMA* 276: 496-502.
Ratcliffe, Matthew. 2017. “Empathy and Psychiatric Illness.” In *The Routledge Handbook of Philosophy of Empathy*. Edited by Heidi L. Maibom. New York: Routledge. Page references at https://pdfs.semanticscholar.org/4e19/7ab6fc7067ac365584a6da5a3498b5275277.pdf
Sadler, John. 2008. “The Instrument Metaphor, Hyponarrativity, and the Generic Clinican.” In *Philosophical Perspectives on Technology and Psychiatry.* Edited by James Phillips. Oxford: Oxford University Press. 23-33.
Searle, John. 1984. *Minds, Brains, and Science*. Cambridge: Harvard University Press. Chapter 2.
Somers, Margaret R. 1994. “The Narrative Constitution of Identity: A Relational and Network Approach.” *Theory and Society*. 23(5): 605-649.
Sophocles. 1957. Sophocles II. Translated by David Grene and Richmond Lattimore. Chicago: University of Chicago Press. 12-13.
Strawson, Galen. 2004. “Against Narrativity.” *Ratio*. 17(4): 428-452.
Sulzer, Sandra H., Noah Weeth Feinstein, and Claire Wendland. 2016. “Assessing Empathy Development in Medical Education: A Systematic Review.” *Medical Education*. 50(3): 300-310.
Tai-Seale, Ming, Thomas McGuire, Weimin Zhang. 2007. “Time Allocation in Primary Care Office Visits.” *Health Services Research*. 42(5): 1871-1894.
Woods, Angela. 2013. “Beyond the wounded storyteller: rethinking narrativity, illness, and embodied self-experience.” In *Health, Illness, and Disease*. Edited by Havi Carel and Rachel Cooper. 113-128.
Woolf, Virginia. 1926/2002. *On Being Ill*. Ashfield: Paris Press.