Let me begin by carving out a domain within, or perhaps adjacent to, narrative complexity. Following Seymour Chatman (1978) and a long-established tradition in literature and related arts, I start with the distinction between story and discourse. The story itself is the narrative. The discourse is in the utterances and sentences that an author uses to tell a story. More particularly, the style of how the story is told is the narration. Since the fictional form I wish to consider is that of movies, I’ll follow David Bordwell (1985, xi) and consider movie narration as “the activity of selecting, arranging, and ordering story material in order to achieve specific, time-bound effects on the perceiver.”

Narration can be simple or complex, and it can contrast with its accompanying narrative. For example, it is easy to imagine a simple story told in a complex manner. Consider the movie Memento (Nolan 2000). It has two interwoven narrational threads—one in color, moving scene by scene backward in time, and one in black and white simultaneously moving forwards. A recut and remix of the film with both threads moving forward in tandem reveals a facile narrative and an uninteresting movie.1 It is also possible to imagine a complex story told in a simple style. Consider a network movie like Valentine’s Day (Marshall 2010). It provides the stories of twenty mostly unrelated characters interwoven by their interactions across a single day, and yet it is told in a linear style—no flashbacks, no apparent jumps in time, no complicated turns. But how does one determine what is simple or complex?

It’s All about Counting

The word simple often has negative connotations and implications—such as simple-minded and simplistic—but this is not true in many academ-
ic disciplines. Simplicity has played a central and positive role in discussions within philosophy, psychology, and computer science and has done so for millennia, for more than a century, and for at least forty years, respectively. Unsurprisingly, the accounts of simplicity in philosophy stem from Aristotle, but they are also found in the work of a number of medieval Scholastics. Much discussion has centered on William of Ockham, the English Franciscan friar of the fourteenth century. Ockham’s dictum (typically called Occam’s Razor) is often expressed as a rubric for doing and developing a science (Carey 2010), but it can easily be expanded to the humanities.

Nonetheless, application of Occam’s Razor is not without problems. It is used to address research situations as something like “among competing theories one should prefer that with the fewest assumptions”—ergo, the simplest. To follow the dictum one should count assumptions before any choice of theory is made. Such enumerations are not always easy. Assumptions, like many things, are often hidden or overlooked, and themselves differentially simple or complex. Thus, except in restricted circumstances (see Goodman 1955; Kemeny 1955), application of Ockham’s tenet seldom progresses beyond a pragmatic rule. Despite this, it still carries considerable weight in scientific discourse.

In my home professional subfield, perceptual psychology, interest in simplicity became popular at the end of the nineteenth century. Ernst Mach ([1897] 1914), the German physicist, first promoted the idea of a perceptual economics. The idea came to full flower with the Gestalt psychologists (see Koffka 1935) under the rubric of prägnanz (German roughly for “pithiness”). Consider perceivers looking at a complex visual array. It is said that, among the many possibilities, they will see the simplest interpretation of what is optically projected to their eyes. Thus, for example, when looking out at an architectural space, they will tend to regularize it, seeing (and assuming that) all walls, ceilings, and floors intersect at right angles even when, as is common in Europe, those joins are not so regular. The rectangular version is thought simpler. Why?

The answer concerns a measure of simplicity. In computer science this metric generalizes from algorithmic-information theory. Roughly, if one were to write an equation or a computer program, or even draw a set of architectural plans to represent a structure, these would be shorter in length, or cleaner in blueprint, for the rectangular building than
for one less regular. The reason is that every intersection can be assumed to be the same in the former; in the latter it cannot. Following Ockham and Kurt Koffka (1935) the simplicity in rectangular form is that it has a single assumption about all intersections.

Wendell Garner (1970), one of my mentors, once wrote that “good patterns have few alternatives.” This means that simple patterns have redundancies and duplications of structure that limit the possibilities of interpreting them. A square is a square, or perhaps a tilted diamond, but it will not be seen as an irregular parallelepiped, a tree, a mountain, or a bird. In contrast, a Rorschach blot may have many possible readings (although these interpretations may shed little light on mental issues; see Lilienfeld, Wood, and Garb 2000).

The obverse of simplicity is, of course, complexity—a word with a striking etymology. Com means “with” and plex means “woven,” so complex things have parts woven together. This is the essence of nonlinearity in science and of interaction in statistics. The parts of a whole are mutually interdependent. Complex can be contrasted with complicated, where plic means folded, and implies that parts of things are hidden from view, folded over but not interwoven. Thus, complicated things can be hard to understand because some parts are hidden. Complex things might be also but because isolating each part does not necessarily bring one closer to understanding the whole. Apart, warp and weft are not at all like the fabric they make together. Over the past thirty years or so it has become common to go from discussions of complexity to discussions of chaos (Çambel 1993; Lewin 1992), but I see no need to go there in this context. My discussion is of numbers that are not very large.

Back to perceptual psychology, Gary Hatfield and William Epstein (1985) proposed that there was not just one kind of simplicity but three and by extension a corresponding number of complexities. The focus in philosophy and computer science has been, regularizing Hatfield and Epstein’s terms a bit, descriptive simplicity—axioms, equations, and programs to describe entities that could be compared. A major focus of the Gestalt psychologists was phenomenal simplicity, the apparent perceptual purity of a given form, which might (or might not) also be captured by a description. The third entry in this collection is process simplicity, the psychological effort needed to run the descriptive program. A short computer program, for example, might have a
loop within it that must execute a million times and take a long time to run, whereas a longer program might accomplish the same task without the loop and calculate the result in less time. Accepting this elaboration of Hatfield and Epstein’s threefold distinction, I believe that cognitive processes should be a critical focus for ideas like narrative and narrational complexity.

**Complexity and Embedding**

In this cross-disciplinary gambit consider next the concept of linguistic complexity. This idea divides a number of ways. Two concern the complexity of a language and the complexity of a sentence or utterance. The former addresses the learnability of a second language given a first. The increased complexity of the second language over the first should make it harder to learn, and perhaps the increased complexity of the first language would make many other second languages easier to learn. In addition, pidgins and creoles that descend from the contact of more standard languages appear to have devolved into linguistic entities simpler than their parent forms (Kortmann and Szmrecsanyi 2012). One approach to their assessment is similar to that in computer science with algorithmic-information theory: describe the abstract logic of the language and count the kernel operations.

Another construal of linguistic complexity addresses the depth of recursion in the application of rules. Recursion wraps the forms of sentence kernels together in a single sentence. Typically, the two rules discussed in this context are branching and embedding. Among the complex sentences in English, most are right-branching structures (about 75 percent). Center-embedded structures are much less common (about 10 percent). Moreover, in polysynthetic languages (those with long words and many morphemes) the relative proportion of the latter is even smaller (Karlsson 2007; Mithin 1984). But linguistic rules provide little rationale for why center embedding is less prevalent. The reason is, in part, that it is psychologically harder to unpack. This is process complexity.

Consider sentences 1 and 2 below. The first is right branching, the second center embedded.

1. The man drove the car that hit the dog that died.
2. The dog the car the man drove hit died.
Sentence 1 seems reasonably comprehensible as clauses dribble off the right end. Its mate, although entirely grammatical, seems like near nonsense to the English reader or listener, particularly with its string of verbs at the end. It is complex in a particular way that makes it less comprehensible. One level of such center embedding is about all we can handle; two and more levels of right branching are no problem.

Is embedding a good model for measuring process complexity? Likely not. Consider the everyday discourse between two people. Center embedding can occur when one person (A) introduces a topic, the other (B) responds, and sometime later A introduces a second topic before fully responding to B. Then that topic continues, but one conversant introduces a third. Eventually, A and B finish up, go back to the second topic, and then return to the first. Stephen Levinson (2013) presented the case for fairly deep center embedding in natural discourse and analyzed examples from recorded conversations between customers and store clerks. He suggested that it is not unusual for such pragmatic center embedding to occur at three, and perhaps up to six and more, levels deep. The importance of this type of structure in discourse, as opposed to that in sentences, is that conversational complexity seems much closer to narrative complexity than does a single complex sentence. Moreover, the people engaged in the discourse show no evidence of any difficulty in processing such embedding. Three and four embeddings seem no more difficult than two, and phenomenally (to the conversants) such conversations seem natural and straightforward.

One can also apply the notion of embedding to many movies. They occur with every flashback and its return, the flash forward. Flashbacks reveal parts of a backstory nested within in the telling of a main story. Moreover, as with sentences and discourse, movies can have multiple center embeddings—flashbacks within flashbacks. Indeed, Bordwell (2011) discussed two movies that are particularly heroic in their nested flashbacking—*Passage to Marseille* (Curtiz 1944) and *The Locket* (Brahm 1944). Each has a flashback within a flashback within a flashback within the main story.

Again, one might look to the depth of flashbacks as one measure of process complexity for popular movies, but there are two problems with this idea. First, nested flashbacks are not common, and all movies without any flashbacks might then be regarded as equally com-
plex. Second, it is not clear that *Passage to Marseille* and *The Locket* are any more difficult to understand than other, more linear movies. Flashbacks simply allow past information to inform the present narrative at the time when the filmmakers want that information to be revealed. This is done all of the time in movies without flashbacks. For example, in *Psycho* (1960) Alfred Hitchcock reveals that Norman Bates’s mother is long dead well after we think we have recently seen her. The embedding of flashbacks may indeed be limited to four levels in popular movies, but this may be because of external constraints. One such limitation is fitting the narrative within the 110 minutes or so of the average film’s length.

Finally, with respect to nested embeddings, Lisa Zunshine (2006, 2007) may be on to something in her outlines of a cognitive-literary theory. The core idea centers on embedded mental states. “I like Kona coffee” reveals one mental state—obviously, one of the speaker’s preferences. “My partner used to think that I didn’t like Kona coffee” reveals two mental states—thinking about and not liking, nested across two people. “I realized that my partner used to think I didn’t like Kona coffee” reveals three nested mental states—[realizing [thinking [not liking]]]—that crisscross people. And, at least logically, this can go on and on.

Zunshine suggests that, without an indication of mental state, prose is boring: “He went to the pharmacy to buy aspirin.” But “He hesitated to go to the pharmacy to buy aspirin” is a bit more interesting. The reader learns about the character’s mental state at a particular time and might wonder why he is in that state. Zunshine rightly couples the analysis to the phenomenon of *theory of mind* in psychology, the notion of attributing beliefs, desires, and intentions to another person, animal, or thing. She goes farther and claims that learning about mental states is the essence of why we like and read literature; indeed, it is the core of social life. Minds fascinate us. We want to know about, and we certainly learn from, the mental states of others. Moreover, as the nesting of mental states gets deeper—“he knows that she knows” (two levels) and “she knows that he knows that she knows” (three levels)—things can get really interesting. Zunshine (2007) claimed that this kind of three-level mental play didn’t appear in English literature until the time of Jane Austen.

Does literature go beyond three levels of embedded mental states? I have no strong intuitions here, but my guess is that when only two people

*Simplicity, Complexity, and Narration* 205
are involved nothing much of interest happens beyond “she knows that he knows that she knows.” Who cares if, further, “he knows” all of that? However, with three or more people involved, perhaps it can. Network novels and network movies (those with many interacting characters) such as Grand Hotel (Goulding 1932) would seem to provide fodder for this possibility and for its assessment.

One caveat here, however, concerns the Dunbar number. This number suggests that these levels may barely tax the human mind. Robin Dunbar (1992) promoted the idea that the size of a living group for a given animal species is limited. The group limit is set in the wild by the constraint that everyone should know everyone else in the group and generally know what everyone thinks of everyone else. Failing these tenets, it is not a stable group and will likely break apart. For human beings this number is claimed to be about 150, although its absolute size has no real importance here. Given its general size and our everyday living, however, knowing what person A thinks of B thinks of C thinks of D seems less impressive as a measure of process complexity.

Interim Summary
Ultimately, whether one is a philosopher, psychologist, computer scientist, linguist, film scholar, literary scholar, or anthropologist, the measurements of simplicity and complexity reduce to counting. What one counts varies across disciplines, but most everyone agrees that three of something is more complex than two of the same thing, which in turn is more complex than one. All else is a variation on this theme. But what should we count? The degree of nesting, center embedded or not, seems narrow and insufficiently fruitful in this regard. Moreover, as we count, shouldn’t we ask whether the measure of complexity has psychological implications for processing? Such implications could be measured as an aversion to complexity or even, as Zunshine suggests, an attraction to it.

Text Complexity and Cognitive Effort
One goal of this chapter is to place the notion of complexity in a cross-disciplinary context. It now seems prudent to come closer to narrative complexity and to discuss reading and writing as a communicative form. In efforts to systematize the craft of writing, style manuals in English
almost universally suggest some form of simplicity. For example, rule 17 of William Strunk and E. B. White (1959) famously suggested, “Omit needless words.” Why?

In part this is a matter of aesthetics and style, like pulling weeds from a garden. But it is also a matter of reducing cognitive difficulty in reading—the amount of effort needed to process text. We build an idea as we scan across the page. The harder the text is to read, the harder it will be to hold on to all parts of this idea before we’ve reached the end. Again, measuring this difficulty reduces to counting. Words are one of the countable units.

But omitting needless words from sentences is only part of the trick. Long sentences are long mostly because they have multiple clauses. Indeed, a popular online dictionary suggests that the very “definition of a complex sentence is the joining of an independent clause and one or more dependent clauses.” This too is about counting—in this case the thought kernels in complex sentences. There is also a third variable, typically discussed as word length, determined by counting number of letters in a word. This variable is highly correlated with two other important factors. The first is word frequency, a reasonable surrogate for the likelihood that the reader or listener knows the word (Strauss, Grzybek, and Altmann 2007). The second is its number of syllables, which often reflects its ease of pronunciation. Difficulties articulating can deflect cognitive resources away from understanding and retaining the meaning of the sentence while coursing through it (Just and Carpenter 1987).

Together sentence length, orthographic word length, and syllable length are the major metrics that have been used to measure textual complexity. The oldest and most widely used is the Flesch-Kincaid Grade Level Index, stemming from the ideas of Rudolf Flesch (1946), then modified and developed for the U.S. Navy (Kincaid et al. 1975). Some U.S. states use this index to reign in the horrible prose of insurance policies and government reports.

The Flesch-Kincaid uses a linear regression formula with two of the variables previously discussed—the average number of words per sentence and the average number of syllables per word. It then norms these values to the reading materials found in the various grade levels of U.S. schools. Consider the examples in sentences 3 and 4:

*Simplicity, Complexity, and Narration* 207
3. This sentence has a Flesch-Kincaid index at the sixth-grade reading level.

4. Because it penalizes polysyllabic words, the index for this sentence makes it more appropriate for a twelfth grader.

The Flesch-Kincaid system was designed to handle texts only for grades one through twelve. Sentence 4 doesn’t seem very complex, so longer and more complex sentences would yield inappropriately large “grade” levels. This kind of effect commonly occurs when one pushes a prediction from a regression analysis to go beyond the domain it is intended to serve.

One relatively unnoticed benefit of Flesch-Kincaid is that complexity is not limited to the cardinal numbers—1, 2, 3, and beyond—as embedding is and as an application of algorithmic-information theory would be. Instead, it is a continuous measure, typically rounded to the first decimal. Thus, Anne Sewell’s Black Beauty is deemed appropriate for readers in grade 4.2 (a few months into the fourth grade) and Mark Twain’s The Adventures of Tom Sawyer for readers in grade 4.5. One can quibble with such designations, but continuous measures are very helpful, allowing consideration of gradations among texts.

The Flesch-Kincaid system is one among many. Clearly such indexes are simplistic and easy to critique (Bailin and Grafstein 2001), even lampoon. Some short monosyllabic words are arcane, some long ones easy to comprehend. Some compound sentences are more understandable than others. “Twas brillig, and the slithy toves did gyre and gimble in the wabe” would supposedly be appropriate for the fourth grade, and the whole of “Jabberwocky” (Carroll 1866), with its one-word exclamation sentences, is “fit” for the first grade. Obviously, misapplication of Flesch-Kincaid could lead one seriously astray.

Nonetheless, I defend such indexes, at least in part. The complexity of literature can be a fuzzy notion, and, sentence by sentence, there are always exceptions to any generalization. A statistical eye toward the norm is necessary in this context, along with the knowledge that these indexes can never be applied meaningfully to single sentences or short texts. Over a larger corpus, such as a novel, the force of particular criticisms is diluted and becomes less important. Anomalies and exceptions across the average flow of prose get statistically minimized. For example, on the
basis of Flesch-Kincaid counts, the following assignments of novels to reading grade levels would seem to have some face validity:

Sixth grade: L. Frank Baum, *The Emerald City of Oz*; Jules Verne, *The Journey to the Center of the Earth*.

And for comparison purposes, Flesch-Kincaid would rate the text of this chapter at grade level 11.1. My mean sentence length is 18.4 words with a mean of 1.66 syllables per word. I don’t have the analogous data for the novels listed here, but, to make things a bit more concrete, those judged to be appropriate for the twelfth grade could have an average sentence length of 21.0 words and 1.65 syllables per word. A novel appropriate for the eighth grade could have an average sentence length of 15.0 words and 1.50 syllables per word. Thus, as one would expect, there are incrementally more “big” words in the high school senior texts than in those suited for middle school. There are also longer sentences, indirectly implying relatively more subordinate clauses.

**Interim Summary**

My main point is that such measures of textual complexity really do have process implications. It seems reasonably uncontroversial that, in terms of readability and comprehensibility, *The Journey to the Center of the Earth* is less complex than *Northanger Abbey* and that *Oliver Twist* and *Wuthering Heights* might be well spaced in between. And yes, as Zunshine’s analysis might suggest, Austen’s story seems narratively more complex than Dickens’s, but none of that kind of complexity is measured by Flesch-Kincaid, only the lengths of words and sentences, which are the tools of narration. It is as if we might trust authors to mold the difficulty of their words and sentences to the overall difficulty of their narrative.

But the main point is that, in daily prose and in literature, we like a certain amount of complexity, particularly as we mature. The overly
complex is too difficult and may not repay the effort to understand, and
the overly simple is vapid. In general, we have an intermediate zone that
we prefer, which of course varies with context and across individuals.

Goldilocks Narration?
When it comes to the popular arts surely there is some zone of compro-
mise between extremes of simple and complex. Borrowing from Robert
Southeby’s nineteenth-century popularization, we might call this the
Goldilocks zone—the porridge is the right temperature, the chair the
right height, the bed the right softness, and the story grabs the read-
er/listener/viewer in an all-encompassing way. The Flesch-Kincaid in-
dex hopes to measure that zone for classroom texts and to provide a
metric useful for teachers and librarians in guiding their students. It is
about choosing an appropriate narrational form as measured indirectly
in words and sentences and hoping the narrative itself will ignite inter-
est. If we can count the appropriate things in the narration, perhaps we
can discover a format that readers, listeners, and art consumers appreci-
ate most, on the assumption that it makes them do a gratifying amount
of cognitive work.

Psychological interest in aesthetics has typically linked interest and
appreciation with physiological state. Wilhelm Wundt may have been the
first to suggest this idea in the nineteenth century, but Daniel Berlyne
(1971) embraced it and made it a central principle in the psychology of
aesthetics. Wundt and Berlyne considered this as a rising and then fall-
ing curve, much like a bell-shaped, or normal, distribution. Hedonic
value—how much we like things—is compared to physiological arous-
al. Consider three cases: When we are aesthetically indifferent to some-
thing our physiological state is unexceptional. Call it normal. When we
like something, our interest increases, and we become more aroused—
our eyes may dilate, our heart may speed, our breathing slow—but only
up to a point. If we are too aroused, we find that state discomforting,
and we must back off; the hedonic value decreases and may even become
aversive. A certain amount of psychological data supports this view, but
there is also a variant more pertinent to this discussion.

Psychology suggests that we as perceivers, like Goldilocks, tend to like
best those things that are dimensionally just right—in this case, neither
too familiar or unfamiliar and neither too simple or complex. To be sure,
there are doubts about this idea (Martindale, Moore, and Borkum 1990), and many of us accept that context and individuality matter greatly (see, for example, Carney, Wlodarski, and Dunbar 2014; Eden, Daalmans, and Johnson 2017). Moreover, the central problem with its corroboration has been any reasonable measurement of simplicity and complexity. Again, this essentially reduces to counting, and some ingenious methods have been used.¹⁰ I won’t copy them here. Instead and more simply, I consider the nature of scenes in popular movies. And, of course, I also count.

The Increasing Narrational Complexity in Popular Movies

Why study popular movies? My students and I have nothing against art film, but we wished to investigate the structure of movies that large numbers of people have viewed. This criterion ensures that many viewers have chosen to watch them as entertainment and have generally enjoyed them. Such movies are thus deemed to tell sufficiently engaging narratives that keep viewers coming back to the medium. Indeed, the British Film Institute (2012) estimated that the average person in the United Kingdom sees more than eighty films per year across all platforms. It seems unlikely that U.S. citizens would be much different.

Which movies? We have chosen to study English-language movies. Most were made in the United States, although the notion of the “Hollywood movie” has been globalized so that many such movies now have some production aspect done in Europe, Asia, or elsewhere. As with art films, we have nothing against non-English-language movies. We acknowledge that there may well be interesting variations in other national corpora, and we would embrace their study, but Hollywood movies seemed like a good place to start.

We also hoped to ensure there was little bias in our choice of films. Thus, instead of selecting them by our own quirks and preferences, we went to online sources and selected those that were among the highest-grossing movies of their particular release years.¹¹ We also hoped to trace historical changes. Thus, we sampled movies released between 1915 and 2015, ten per year at five-year intervals, across many genres—mostly dramas, comedies, action films, adventure films, and animations, and we have analyzed these movies in a large number of ways (see Cutting 2016a, 2016b; Cutting, DeLong, and Brunick 2018, for overviews).

Popular movies have many cuts—abrupt breaks between the separate
shots—but shots are not my unit of interest here. Like plays, movies can also be divided into scenes, each of which typically has seven to a dozen shots, although the range can be from one to over one hundred. A scene boundary in theater is generally defined by a change in location, a change in characters, or a change in time (Polking 1990), and the same holds true for movies. I call all combinations of these changes narrational shifts. To assess their occurrence, I focus on a twenty-four-movie subset of our overall sample—one drama, one comedy, and one action film released at ten-year intervals from 1940 to 2010. Importantly, all of these movies are roughly the same duration—about 115 minutes.

My students and I determined narrational shifts in two ways. First, we had undergraduate viewers watch each movie twice and, on the second pass, mark the frame number of the first shot in every new scene. We let them define a scene in any way they chose. Second, we duplicated their task more intensively, going through these same movies frame by frame and counting every change in location, characters, or time. Correspondence between our frame numbers and those of the undergraduates was good. To be sure, not every student segmented the movies in the same way, but almost every location, character, or time change we marked was also noted as a scene boundary by at least one of the three students who segmented each film (Cutting, Brunick, and Candan 2012; Cutting 2014). My interest here is not in the separate types of shifts but simply in how many there are, regardless of whether all three dimensions shifted at the same time, only two, or just one. Neurophysiological evidence suggests that the brains of movie viewers respond similarly to narrational shifts across cuts and differently to the within-scene camera changes across cuts (Magliano and Zacks 2011).

One caveat is important here. These shifts mark boundaries in scenes, but they also separate what I call subscenes. Subscenes are parts of larger units called sequences, of which almost every movie has at least one. One prototype is the chase sequence found in many action films, but they are common in other genres as well. Consider here the launch sequence in The Martian (Scott 2015) near the beginning of the film’s climax. In a span of four minutes, the movie brings together seven different locations—the surface of Mars, the rescue ship orbiting Mars, the inside of NASA/Houston Control, the outside of NASA headquarters, and public arenas in New York City, London, and Beijing. It cuts
among these locations thirty-five times. Each location has different characters, but all subscenes follow a linear and contiguous time frame, which binds the seven narrative threads together into a single sequence of thirty-six subscenes.  

Now consider the data of the twenty-four movies. As can be seen in figure 19, there has been a general rise in the number of narrational shifts across scenes and subscenes. They more than doubled between 1940 and 2010. This rise, I would argue, suggests that there has been increasing narrational complexity in popular movies during this time. That is, movies jump around among locations, characters, and time frames considerably more often now than they used to. Since the movies within and across these release years are about the same length, this means that scenes and subscenes have gotten shorter. In 1940 they averaged about

Fig. 19. A representation of the number of narrational shifts (changes in location, characters, time, or their combination) in twenty-four movies (shown as dots) released from 1940 through 2010. Created by the author.
ninety seconds long but by 2010 they were only about thirty seconds in duration, although, as implied by The Martian’s launch sequence, there is great variability in length.

Other factors are worth considering. The number of shots per film has also shown a marked but gradual increase across these same years, from just under five hundred to over two thousand per movie, and the corresponding mean shot duration has necessarily decreased—from a bit over twelve to just under four seconds. Why has this occurred? Almost 99 percent of all shots in movies since the 1980s are separated by cuts, and each cut will typically dictate an eye movement toward the central content of the next shot (Smith 2012). This eye movement reflects the attention being paid to the movie by the viewer. Shot changes and scene changes in contemporary movies keep our minds busier and perhaps increase our engagement in the narrative. Again, elaborating from Hatfield and Epstein (1985), this process complexity may be something that we enjoy.

What has caused this rising number of narrational shifts in popular movies? I don’t really know, but I suspect it revolves around a decades-long interplay of filmmakers, technological change, and audience knowledge. On the one hand, film editors often decry the use of too many rapidly cut sequences (Longwell 2007), but they seem to produce them anyway. Is there a way to sift through these data and find evidence that audiences contribute to this trend? The data in figure 20 suggest that there might be.

Each of these twenty-four movies was very successful. But of the three movies in the sample released each year, I first ranked them according to their reported gross earnings—a direct measure of the number of viewers who saw the films. I then considered separately the most successful of the triplet of movies in each release year and those that came in second and third. Shown across the three panels are results, plotted by release year and their number of narrational shifts. Notice that the highest-grossing films have a steeper slope. The slope of the second-ranked films is shallower, and that for those of the third rank even more so. Other things assumed being equal—relatively large budgets, general promotion in other media, and roughly equivalent star power—these results, although fairly slim, suggest that viewers have responded positively to complexity. Among the successful movies of any given year, those with a greater number of narrational shifts may have been rewarded with larger audiences. It would seem, then, that insofar as an
increase in the number of narrational shifts might reflect process complexity, viewers like it—at least to the extent currently reached by contemporary movies. One can easily imagine that there must be an upper limit, but it would seem that even Inception (Nolan 2010), represented by the highest point in figures 19 and 20 and with its over four hundred narrative shifts across four dream states and diegetic reality, has not yet reached it. Two final points: First, Inception was clearly designed to be

Simplicity, Complexity, and Narration 215
seen more than once, as if the narrative cannot be completely grasped in a single viewing. But, second, the data suggest that audiences may have grown in the amount of complexity in the narratives that they wish to consume. Perhaps they are intellectually driven to this zone between the overly complex and the overly simple.

**Complexity in Numbers**

Complexity is defined in many ways across many disciplines. The term’s etymology—roughly “interwoven”—suggests that its unpacking can rarely be straightforward. Whatever makes for complexity has parts that may interact in unpredictable ways. Elaborating a notion from Hatfield and Epstein (1985), I find that narrative complexity may best be thought of in terms of cognitive processing: how hard the mind has to work to understand something. I call this *process complexity*. Fortunately concerning narratives, a certain amount of such cognitive work can be enjoyable, as advocated by Zunshine (2006). Working to unravel complexity—particularly when it deals with our theories of the minds of characters—appears to hold our attention and interest. When mind reading is too facile, it can be boring. Likely when a narrative starts out with too many characters with too little known about how they think, we many lose interest as well. Thus, there may be a narrative sweet spot, a Goldilocks zone, which may also vary greatly across readers and viewers.

Across myriad disciplines the consideration of complexity entails counting, with the general idea that more means harder (to process). Here I counted the number of locations, characters, and time shifts in popular movies from 1940 to 2010 and found a marked increase in the number of such shifts. This, I would contend, makes the various threads of a narrative a bit more difficult to hold on to, but perhaps enjoyably so. At least part of this trend may be driven by viewers; evidence from movie earnings and the prevalence of narrative shifts suggests that the average moviegoer may like these jumps in narration. Why this might be the case is unclear. It may be that movies have “grown” complex to better fit our preexisting cognitive capacities. It may also be that the long-term penetration of movies into popular culture has increased our capacity to follow more complex narration.

Critically, I think these results and similar ones, like those found with the Flesch-Kincaid index—measuring novels, are measures of *narrational*
complexity, not of narrative complexity. I have followed Chatman (1978) and others in separating stories (narratives) from how they are told (narration). Driving this wedge between surface style and underlying form however, leaves me in the awkward position of possibly skirting what may be more important aspects of the complexity of stories. So here is a final thought.

Kristen Thompson has suggested that the “most basic principle of Hollywood cinema is that a narrative should consist of a chain of causes and effects that is easy for the spectator to follow.” But easy is not simple. “The glory of the Hollywood system lies in its ability to allow its finest scriptwriters, directors, and other creators to weave an intricate web of character, event, time, and space that can be transparently obvious” (1999, 10, 11; my emphasis). Simple narratives are often called linear: A causes B causes C and so forth. Intricate narratives fold over causes and effects and often place narrational time between them. Such a gap, in a near parody of the grammatical blunder, creates a “dangling cause,” a foreshadowing. But in popular films, as opposed to art films, all the causes are usually wrapped up in a final effect, where protagonists have achieved their goal, often with an epilogue tacked on for the restoration of normality.

The “intricate web of character, event, time, and space” seems a fruitful venue for measuring narrative complexity, and of course I have already discussed changes in characters, time, and space (or place) as the rationale for segmenting events (scenes and subscenes) from one another. Narrative complexity would seem to lie in the nature of this “intricate web.” One might draw this web as a network over narrational time (beginning to end), enumerating the density of nodes (characters and their goals at particular times and places), the number of edges (threads or lines connecting those nodes), and the length of those edges (the length of time over which they dangle). Parallel action, flashbacks and dreams, false leads, and other storytelling devices would contribute to the intricacy of this web. Hopefully, the combined measure of these—achieved by some sort of counting—would also reflect both psychological process and narrative complexity.

Notes
1. For a discussion of the reediting and rewatching of Memento see Morrow (2103).
2. A counter to the simplicity principle, also called the minimum principle, is the likelihood principle. That is, we perceive what is most likely, not what is
simplest. Likelihood would require a Bayesian analysis of prior experience. See Peter van der Helm (2000) for an analysis of the debate.

3. Gregory Chaitin’s (1966) and Andrey Kolmogorov’s (1968) works are considered founding documents of algorithmic-information theory. This idea is also sometimes called “minimum description length” (Rissanen, 1978).

4. The Mandelbrot (1982) set, which generates a fractal design of immense (even infinite) complexity, can be written in an equation with three variables.

5. Discussions of theory of mind began with David Premack and Guy Woodruff’s (1978) experiments with a chimpanzee.

6. The full context of Dunbar’s (1992) claim is that group size in primates is limited by brain size. This is not necessarily true, but we think that claim is independent of the point we wish to make.


8. See “Flesch-Kincaid Grade Level” (n.d.) for an analysis of standard texts in U.S. schools and their reading level.

9. There is a convenient online tool that counts words, syllables, sentences, and many other things; see Rocca (2018).

10. See Robert McCall and Paul McGhee (1977) for an early discussion of the discrepancy hypothesis, which compares stimulus familiarity with interest by infants, and Celeste Kidd, Steven Piantadosi, and Richard Aslin (2012) for a rigorous, information-theoretic analysis of complexity and infant interest that makes good on the claims of the discrepancy hypothesis.

11. The sites we used were Box Office Mojo (www.boxofficemojo.com) and the Internet Movie Database (www.imdb.com), more or less continuously between 2009 and 2016.

12. The time frame is contiguous when allowing for speed-of-light differences in reactions between Earth and Mars.

13. As shown in figure 19, the increase in narrational shifts in movies over release years from 1940 to 2010 is statistically reliable, $t(22) = 4.37, p < .0002, d = 1.86$. The dark line is the best-fitting statistical trend (regression line) that minimizes the sum of the (squared) vertical distances from it to the twenty-four points. The darker-shaded area represents the statistical confidence one should have in that trend. This is called the 95 percent confidence interval; that is, given this sample of movies one can be confident that the trend for all movies of this era would, 95 percent of the time, lie within the bounds of this region. The lighter-shaded region is the 95 percent confidence interval on the data of this sample, given the slope of regression line; that is, 95 percent of these data lie between these bounds, and, given twenty-four points, they all do, with one point very close to the upper line. The vertical axis is logarithmically scaled.
14. It is often suggested in public media that movies have shortened their shots because people have shorter attention spans than they used to. There is no evidence for this claim, and, given the evidence for rising spatial abilities worldwide (Flynn 2013), the reverse might be true.

15. The results shown in figure 20 are also statistically reliable: release year by rank within release year, \( t(20) = 2.65, p < .015, d = 1.18 \). In each panel the 95 percent confidence interval on the fitted lines is shown. These data suggest that more viewers see movies with more narrational shifts. We suggest that this means that viewers are more engaged when there are many such shifts. Again, the vertical axis is a log scale.

References


Simplicity, Complexity, and Narration 219


*Simplicity, Complexity, and Narration* 221


Scott, Ridley, dir. 2015. The Martian. Los Angeles: Twentieth Century Fox Home Entertainment. DVD.


