A|B|A|I
Association for Behavior Analysis International

COMMENTARY

CrossMark

An Emotional Appeal for the Development of Empirical Research on Narrative

Thomas S. Critchfield[1]

There must be something powerful in a good story, because human beings routinely tell them to entertain, persuade, and convey rules or other important information. Hineline (2018) therefore proposed making the analysis of narrative a priority,[1] for two reasons. First, although a general-purpose science of behavior should examine things that members of society deem to be important, behavior scientists have not launched systematic inquiries into the mechanics of storytelling. Second, an understanding of narrative might support the strategic use of storytelling to advance behavior science and the profession of behavior analysis with the general public. Regardless of the goal, the underlying challenge lies in determining *what* is powerful in a good story.

Following Grant (2005), Hineline (2018) proposed that narrative is a process of manipulating motivating operations (using different terminology, Barnes-Holmes and Barnes-Holmes (2002) argued something similar). When a speaker's[2] story is structured and sequenced properly, he maintained, people will attend persistently and enthusiastically. As an example of this effect, Hineline described his own barely controlled habit of reading mystery novels late into the night in order to learn how plot points would resolve. In Hineline's account, the particulars of a story—the sequencing of events that creates uncertainty, defines a mystery, promotes ambiguity, and so forth—contribute mightily to listener engagement. In essence, what the listener learns early on makes it reinforcing to find out more later.[3]

---

[1]It's reasonable to assume that we in the behavioral psychology community are not masters of the narrative arts. This is nothing new among scientists. More than a century ago, Galton (1908) bemoaned scientists' woeful narrative repertoires: "The expressions used in them were so obscure, the grammar so bad, and the arrangement so faulty, that they were scarcely intelligible on a first reading. The writers . . . may have been successful investigators but their powers of literary exposition were of a sadly low order" (p. 118).

[2]For consistency throughout this article I will employ Skinner's generic terms of *speaker* and *listener*, even though in some cases more specific terms, like *writer* and *reader,* would be appropriate.

[3]Both Hineline and Barnes-Holmes and Barnes-Holmes (2002) used the term *coherence* to describe the (speaking loosely) feeling of satisfaction that humans experience when a story resolves.

✉ Thomas S. Critchfield
  tscritc@ilstu.edu

[1] Department of Psychology, Illinois State University, Campus Box 4620, Normal, IL 61790, USA

In Search of Dependent Variables

As a means of illuminating the relevant effects, Hineline (2018) devised a way to diagram plot developments, which can be thought of as verbal independent variables that contribute to a listener's interest. Although Hineline's conceptual analysis is helpful, narrative will be truly understood when we have data to show how it works, and the Hineline account requires further development before it can support empirical analysis. In particular, to properly evaluate independent variables (like plot developments) one needs good *dependent* variables. To wit: when Dr. Hineline reads his novel late into the night, it is clear that *something* important is unfolding during those critical hours—but what? Because *verbal* implies speaker–listener relations (Skinner, 1957), the "something" of interest must be some kind of listener behavior playing out in real time. A good story, according to Hineline, promotes behavioral outcomes that might be described as "sustained listening," "paying attention," and "being engaged with a story," but Hineline did not say how these variables might be operationalized and measured in a research program.

Some inspiration regarding what to measure in studies of narrative might be taken from observers of literature who have suggested that successful stories tend to follow one of several standard patterns (e.g., Abbott, 2005; Booker, 2006; Foster-Harris, 1959; Tobias, 1993). An example can be found in the "story shapes" described by novelist Kurt Vonnegut (e.g., 2005) in a framework that he first created to pitch a master's thesis topic at the University of Chicago.[4] Vonnegut characterized story shapes as reflecting shifts in the well-being of a central character, as in the following examples:

> "Man in a Hole" stories: The main character gets into trouble then gets out of it.

> "Bad to Worse" stories: The main character starts off in bad shape and then things get really bad.

> "Cinderella" stories: The main character starts in dire circumstances, temporarily achieves good fortune, loses it, and ultimately winds up better off than before the loss.

Figure 1 (top panel) shows how Vonnegut (2005) diagrammed one of these story shapes. Good fortune, measured on the vertical axis, fluctuates over time, which proceeds across the horizontal axis. Vonnegut's shapes summarize the surface features of a story, that is, the tenor of events affecting a central character.

Verbal behavior (like that in a good story) exerts its functional effects on listeners, who of course don't personally experience what a story's character does. They may, however, emit strong empathetic responses (Grant, 2005). This can happen via derived stimulus relations in which the listener experiences a sort of equivalence between himself and the main character, or via extended stimulus relations that make a story's fictional events "feel real" (Barnes-Holmes & Barnes-Holmes, 2002; Hineline, 2018).

---

[4] Although the university rejected this idea "because it was so simple, and looked like too much fun" (Vonnegut's interpretation, anyway; quoted in Gonzalez, 2014), Vonnegut (e.g., 2005) continued to promote it throughout his literary career.

To the extent that a listener's joy and suffering covary with those of the main character, a Vonnegut story shape becomes, in effect, an account of fluctuations in listener emotion (Fig. 1, bottom panel). Call this an "emotional arc" (Reagan, Mitchell, Kiley, Danforth, & Dodds, 2016, p. 1), and imagine it as the dependent variable in a study of how plot developments (the independent variables of interest) affect listeners.

Among behavior scientists, *emotion* is a potentially controversial concept, but I invoke it for three reasons—first and foremost because of common sense: anyone who has ever digested a good story knows that the listener experience often involves emotion.[5] Second, a focus on emotion is consistent with Hineline's (2018) conceptual framework and with behavior theory in general. As a preliminary comment on this correspondence, note that both Hineline and Grant (2005) described plot as a manipulation of listener motivating operations, and Skinner (e.g., 1945, 1953) described emotions as one component of fluctuating motivating operations.[6] Third, for practical reasons, in the study of story effects, emotional responses might actually be a more useful dependent variable than certain "public" behaviors. To explain why requires a brief review of the standard behavior-theory account of emotions that Skinner (e.g., 1945, 1953) pioneered.

Emotional Dependent Variables

The Ontological Status of Emotion

Skinner regarded emotions as epiphenomenal, by which he meant that they are responses whose occurrence requires explanation. He did not question their existence or imply that they are unrelated to other behavior. For instance, Estes and Skinner (1941) showed that once emotional responses occur they may compete with other behavior. Skinner (1953) suggested that emotion—including that evoked by verbal behavior—also can potentiate other behavior:

> In a "pep talk" a coach may take advantage of the fact that players exert themselves more aggressively against their opponents if they have been made angry. The skilled cross-examiner may use the same procedure to lead a witness to emit verbal responses which would otherwise be withheld. Soldiers and civilian populations are aroused to aggressive action with stories of atrocities, reminders of present or past injuries, and so on. (pp. 169–170)

Yet Skinner eschewed any necessary role for emotions in an empirical science of behavior because, he maintained, they are collateral effects of the same circumstances that evoke public behavior. For instance, Skinner (1945) asserted that a person's behavior "[d]oes not change because he feels anxious; it changes because of the

---

[5] This may be revealing too much, but I got teary-eyed at the ending of Shirley Jackson's "The Lottery," when first hearing Hamlet's classic soliloquy, and when Disney's Mulan retrieved the arrow that no male soldier was able to reach.

[6] For example, "Any extreme deprivation probably acts as an emotional operation" (Skinner, 1953, p. 165). Note that scientists who specialize in the study of emotion also regard it as closely tied to motivation (e.g., Lang, Bradley, & Cuthbert, 1998).
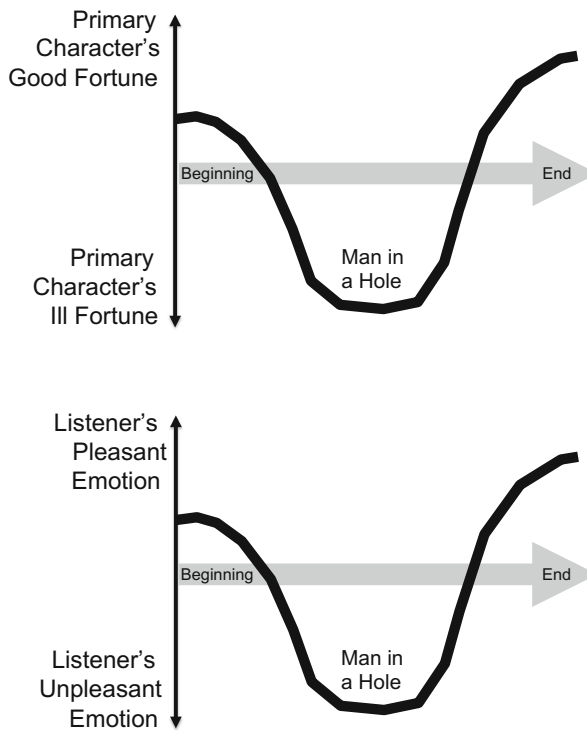
Fig. 1 Top: One of the common "story shapes" described by Kurt Vonnegut. The vertical axis describes variations in the fortunes of a main character, and the horizontal axis charts the progression of the story. Based on the first diagram in Vonnegut (2005). Bottom: The same story shape expressed as fluctuations in the emotional experience of the listener

aversive contingencies that generate the condition felt as anxiety. The change in feeling and the change in behavior have a common cause" (pp. 61–62). From this perspective, emotions have "no functional significance, either in a theoretical analysis or the practical control of behavior" (Skinner, 1953, p. 181) because when one knows the environmental circumstances (historical and contemporaneous) and related public behaviors nothing more is to be learned by attending to emotion. This stance is evident in Skinner's (1953) emphasis on public behaviors in the following passage:

A particularly important emotional predisposition is that in which the individual *favors* a particular person, group, or state of affairs. It is hard to define the particular consequences of "favorable" behavior, but a fairly specific effect can often be discovered. A politician may arrange political rallies, kiss babies, publish favorable autobiographical details, and so on, only to strengthen one very specific response on the part of the electorate—placing a mark on a ballot opposite his name. An author or playwright generates favorable responses toward his characters by depicting them in situations which strengthen such behavior . . . and in this way he increases the chances that his book or play will be "liked"; but the behavior at issue may be nothing more than the purchase of books or tickets or the spreading of favorable reports. Part of the effect here is reinforcement, but we

may also distinguish a kind of operation which must be classed as emotional.
(p. 170; emphasis in original)

Even if Skinner's "not causal" stance is taken at face value, there may be an Achilles' heel to his perspective when it comes to determining how a good story affects a listener. When possible, dependent variables should reveal behavior change in something close to real time (e.g., Johnston & Pennypacker, 1980; Sidman, 1960), whereas the cooccurring public behaviors that Skinner mentioned tend to happen infrequently and/or over extended time frames. For example, a person may buy a given book only once, and may spread favorable reports about that book only a few times across weeks or years, with neither effect occurring at the precise moment when features of the story establish relevant behavioral dispositions.[7] When Dr. Hineline becomes ensconced in a mystery story, we can determine that he stares into his book for 4 h rather than, say, 4 min, but it is difficult to relate this effect to any particular feature of plot. In short, many behaviors that a good story kicks into motion are too coarse to be useful as dependent measures, and attempts to employ them to evaluate insights derived from Hineline's molecular[8] plot diagrams are unlikely to get very far. But given the possible link between emotion and motivating operations, and the common-sense observation that effective storytelling is partly a manipulation of listener emotions on a momentary scale, why not measure emotion in order to gain a preliminary handle on how elements of plot, as described by Hineline (2018), affect listeners?

Justification for Measuring Emotional Responses

Psycholinguists and cognitive scientists have long studied the emotional effects of verbal behavior on listeners, and a key dimension in their analyses is valence, or the general pleasantness versus unpleasantness of listener response (e.g., Boucher & Osgood, 1969). Valence is deemed important in part because it correlates with many interesting behavioral outcomes. When verbal behavior creates pleasant emotion, listeners tend to regard the speaker as familiar, competent, and trustworthy, and they tend to recall and be influenced by the speaker's persuasive messages (Avey, Avolio, & Luthans, 2011; Floh, Kohler, & Zauner, 2013; Garcia-Marques, Mackey, Claypool, & Garcia-Marques, 2004; Norman, Avolio, & Luthans, 2010; Petty, Schumann, Richman, & Strathman, 1993). Figure 2 shows a simple example of this kind of convergence. In an unpublished pilot study, Derek Reed and I asked college students to read 150-word passages (from an article by Rogers and Skinner [1956] on the concept of "control") for which we had determined the mean valence of the component words. The students then made some big-picture ratings about the passages. Mean word valence correlated strongly and positively with ratings of the overall pleasantness of the passages (left

---

[7] Note as well that neither necessarily takes place within an experimenter's scope of observation.
[8] Hineline (2018) might take issue with this characterization. He argued that molar relations unite plot and listener behavior, and about this he is not wrong. After all, the emotional impact of Romeo's death depends on the full series of missteps and misfortunes that precede it. Nevertheless, a rather punctate emotional response accompanies an expository moment like this, and exemplifies the kind of effect that needs to be clearly recorded in studies of narrative.
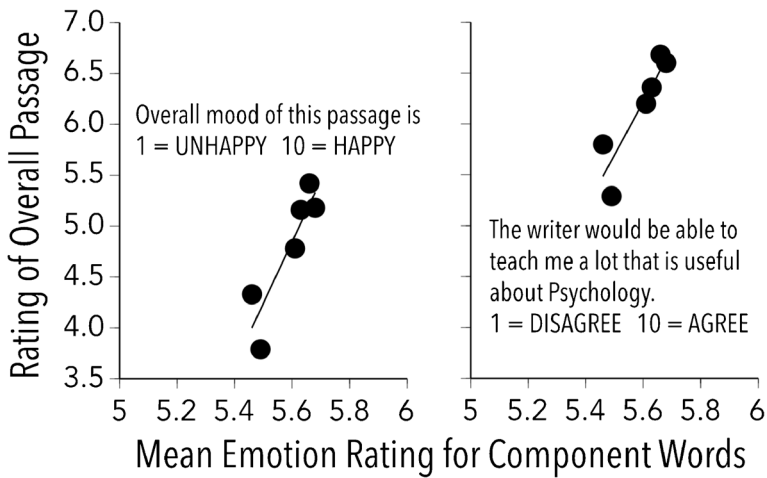
Fig. 2 Correlation between the mean normative emotion ratings (1 = Unpleasant, 9 = Pleasant) of the words contained in brief text passages and overall ratings made by readers of those passages. See text for additional details. Previously unpublished data collected with Derek Reed

panel) and with students' impression that the author would be an informative source for learning more about psychology (right panel). Outcomes like these seem relevant to sustaining a listener's "engagement."

Researchers outside of behavior science have found it convenient and productive to measure emotion by simply asking listeners about it.[9] For example, in studies of the emotional effects of words, the words are presented one at a time to research participants, who rate their gut-level emotional responses on a scale ranging from unpleasant to pleasant (e.g., Boucher & Osgood, 1969; Dodds et al., 2015; Warriner, Kuperman, & Brysbaert, 2013). This procedure has been used to generate quantitative norms of listener emotional response for thousands of words in English and several other languages (e.g., Dodds et al., 2015; Stadthagen-Gonzales, Imbault, Perez Sanchez, & Brysbaert, 2016; Warriner et al., 2013)—these are what Reed and I used to determine the mean valence of passages in the abscissas of Fig. 1. Table 1 shows examples of English words that, normatively speaking, are experienced as especially unpleasant, neutral, or especially pleasant. Or, rather, to be precise, these are words that people *say* they experience in particular ways, because the underlying methodology relies on verbal reports. Yet as noted in the preceding paragraph self-reported emotional responses correlate with useful and interesting behavioral outcomes. Such findings do not prove that self-reports of emotion are veridical, but they do demonstrate how, in many cases, the reports have strong predictive validity (for a detailed discussion of this distinction, see Critchfield, Tucker, & Vuchinich, 1998).

Regarding the use of verbal reports as measurement, it should be noted that applied behavior analysts take much the same approach in social validity assessments (Carr,

---

[9] There are lots of other approaches. This is a multimethod research area (Bradley & Lang, 2000a). It is important to note, however, the psychometric properties of emotion rating scales have been thoroughly explored, and rating scales may perform better than some more elaborate methods (e.g., Bradley & Lang, 1994). Those who are skeptical of measuring emotion should start by digesting the relevant literature, which is substantial.

Table 1 *Some Words that, Normatively Speaking, are Experienced as Especially Unpleasant and Especially Pleasant (from the Corpus Developed by Warriner et al.,* 2013*)*

Strongly Unpleasant	Neutral	Strongly Pleasant
asphyxiation	taxicab	vacation
torture	episode	enjoyment
leukemia	milky	loveable
racism	ajar	sunshine
homicide	principle	laughter
herpes	chalk	prize
abuse	photon	pizza
worthless	questionnaire	hug
suicide	tolerable	tranquility
unhappy	urban	puppy

Austin, Britton, Kellum, & Bailey, 1999; Wolf, 1978). The focus is on how consumers feel about behavioral interventions, and the most expedient way to find out is to ask them. Just as in studies of how verbal behavior evokes emotion, there is no direct corroboration of these listener verbal reports (Carr et al., 1999), but the reports are assumed to correspond to important behavioral outcomes (such as embracing or rejecting interventions).

Despite this parallel, it would be surprising if some readers did not fall back on our discipline's customary admonitions against verbal-report–based measurement—for example, "A subject's description of his own. .. behavior would not normally be accepted as a measure of his actual behavior unless it were independently substantiated" (Baer, Wolf, & Risley, 1968, p. 93). Yet such objections reflect a rather simplistic conception of measurement. For example, the general concern about correspondence with events of interest is not unique to verbal-report measurement, and the specific emphasis on "independent substantiation" is often misplaced, because it implies that the researcher cannot know or arrange the conditions under which good correspondence occurs. For elaboration on these issues see Critchfield et al. (1998).

Friman, Hayes, and Wilson (1998) cast further doubt on standard objections to verbal-report measurement as it applies specifically to private events such as emotion. Drawing upon observations of individuals experiencing anxiety problems, they proposed that the human capacity to form derived stimulus relations functionally entangles emotional responses and verbal statements of those responses.[10] This proposal is consistent with Sidman's (2000) theory of stimulus equivalence in which stimuli, responses, and consequences, when brought together in contingencies, acquire shared functions, as well as with research showing that private events can enter into stimulus

---

[10] Here is a short and incomplete account of a complex issue. Verbal-emotional responses appear to have the same experiential origins as nonverbal emotional responses (a person may experience the autonomic correlates of fear and speak of feeling fearful due to the same terrifying experience). Verbal-emotional responses also can elicit the same autonomic responses as a terrifying experience. They can be potentiated and suppressed by the same circumstances that affect nonverbal emotional behavior. Once evoked, they tend to interrupt ongoing behavior, just as nonverbal emotional responses may. In these ways verbal-emotional responses are functionally equivalent to nonverbal emotional responses and could therefore be measured, and modified, as a proxy to those.

relations (e.g., DeGrandpre, Bickell, & Higgins, 1992). Friman et al. (1998) therefore concluded that

> Reports of anxiety are not necessarily mere statements whose sole function is communication. . . . The reports can also be reactive and thus generate untoward effects. . . . In a colloquial sense, anxiety is what anxiety does, and what it does includes what anxious persons *say* it does. . . . Recounting a highly aversive event often instigates some (often much) of the arousal and avoidance responding that was occasioned by the event itself. (pp. 144–145; emphasis in original)

Put simply, it is a mistake, in the standard conception of measurement, to think of verbal reports as being *about* emotional responses because, in fact, they *are* emotional responses, with functional similarities to their more private counterparts. In this sense, they constitute a suitable focus when the goal is to study emotional responses.

In Vivo Measurement of Emotional Responses to Stories

Those who study literature maintain that the first rule of storytelling is "show, don't tell" (Lubbock, 1921), so I now proceed to illustrate how emotional responses experienced by listeners during the course of a story might be monitored, with an eye toward eventually testing some of the hypotheses about narrative advanced by Hineline (2018). Readers of a certain age may remember when it was popular for television networks, during debates between political candidates, to present an on-screen graphic that showed "emotional" responses of presumably typical voters. Those individuals held a dial that they could turn to indicate how their feelings for a candidate changed in real time as the candidate spoke. A line on the television viewer's screen rose and fell accordingly.[11] Imagine instead a dial controlled by a listener to express emotional valence experienced while consuming one of Hineline's (2018) mystery stories. The dial presumably would create curves similar to those of Fig. 1. By noting where momentary fluctuations in emotion occur, it should be possible to identify specific story elements that correlate with these changes, and thereby to check some of Hineline's assumptions about what makes for a successful story.

To my knowledge, the "emotion dial" has not been employed in studies of narrative, but one research team has begun to explore a conceptually related means of empirically mapping emotional arcs. A brief digression is required to explain their method. Recall that in many studies of the emotional impact of verbal behavior the unit of analysis is the individual word (I described previously how words are rated for general pleasantness vs. unpleasantness). For each of several hundred books, Reagan et al. (2016) estimated the valence of the normative emotional response associated with each successive word. Then for each book they plotted fluctuations in word emotion using something akin to a sliding mean encompassing a small window of words, and followed this measure across the length of the story. Statistical tools were employed to smooth out the resulting curves and to identify similarities in the curves for different

---

[11] Actually, television networks periodically resurrect this procedure, as during the 2016 presidential campaign (Izenson, 2016).
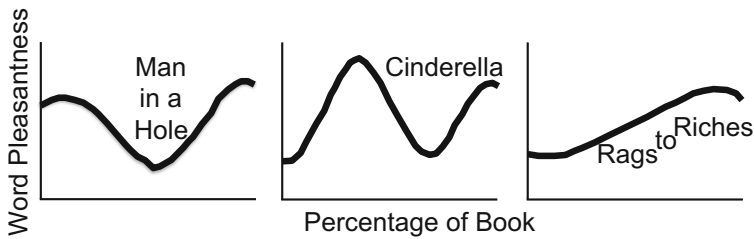
Fig. 3 Empirically-determined emotional story arcs showing how aggregate word pleasantness varies as a story unfolds. Fashioned after Fig. 4 of Reagan, et al. (2016), who derived common story arcs from hundreds of books. These curves resulted after using statistical techniques to smooth the curves of individual books and to determine the features of curves shared by multiple books

books. Reagan et al. found that a surprising number of books employed one of a small handful of global story shapes, including those illustrated in Fig. 3. High-frequency arcs included patterns that Vonnegut had anticipated (e.g., "Man in a Hole" and "Cinderella") plus a few familiar others, like "Rags to Riches," in which an initially unpleasant listener emotional experience gives way to a more pleasant one.

The findings of Reagan et al. (2016) provide a hint of how listener emotion might track the unfolding of a story, which is entirely consistent with Hineline's (2018) motivating-operations account of narrative, because it is reasonable to assume that uncertainty, mystery, and ambiguity may have emotion-evoking properties that render "knowing more" reinforcing. But there are two problems. First, the story arcs of Fig. 3 suggest how listener emotional response might fluctuate, but not what aspects of the story might be responsible for this. Second, Fig. 3 does not actually document in vivo listener behavior; rather, Reagan et al. used normative word-emotion data to project fluctuations in *plausible* listener reactions. Mapping actual listener responses—as per the hypothetical "emotion dial"—would be critical to the needed analysis.

Combining listener-derived emotional story shapes with Hineline's (2018) diagramming conventions would address both problems and thereby provide insights about which specific plot contrivances contribute to which emotional effects. For example, Fig. 4 (top panel) shows the "un-smoothed" emotional arc of the final novel in J. K. Rowling's (2009) *Harry Potter* series. Let us imagine that this was obtained with the "emotion dial" (rather than with the Reagan et al. [2016] procedure) and therefore reflects an actual listener's reactions to the unfolding story. Ebbs and flows of emotion map clearly onto specific plot developments, which include both unpleasant lows (e.g., the torture of Hermione) and pleasant highs (e.g., Harry's reunion with his "adoptive" family, the Weasleys, and a happily-ever-after scene at the end of the book).

Consistent with Hineline's (2018) analysis, it might be expected that listener emotion traces, not to specific events, but rather to *sequences* of events that heighten the story's motivating operations. Casual inspection of Fig. 4 suggests that the lows help to set up the highs and vice versa. No analysis of narrative is complete without an understanding of such interactive effects (e.g., Aristotle, 2015; Bennett & Royle, 2009), on which Hineline's plot diagrams might shed some light. The possible importance of interactive effects can be appreciated by examining Fig. 4 (bottom), which shows a modified *Harry Potter* story arc in which all of the unpleasant lows have all been omitted. In this narrative
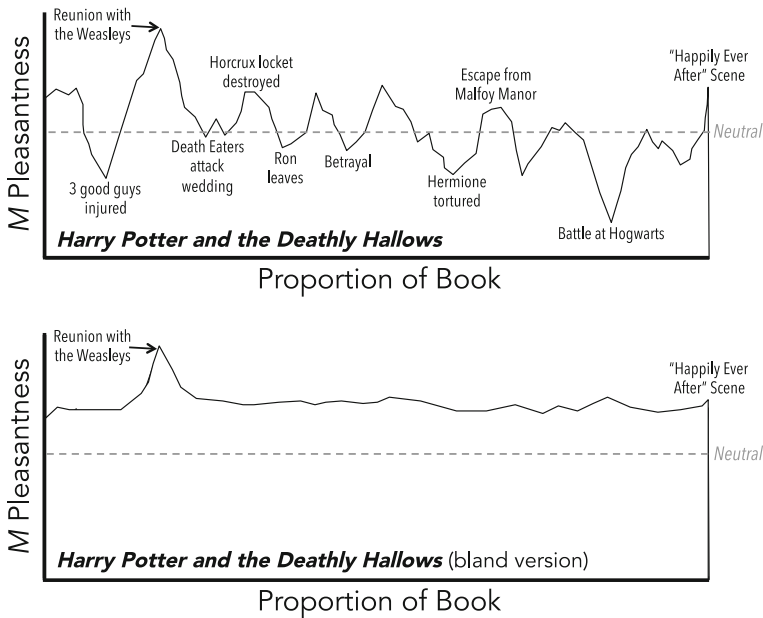
Fig. 4 Top: Empirically-determined emotional story arc of a popular novel, annotated with several key plot developments. Based on Fig. 2 of Reagan, et al. (2016). Bottom: The same story arc imagined with all hints of unpleasantness removed. See text for additional discussion

befitting a Teletubbies®[12] episode, only pleasant things happen, and without the lows it is questionable whether any reader would care about Harry's family reunion or that happily-ever-after scene. The effects that have been stripped away are, I believe, precisely the kind that Hineline's (2018) motivating-operations account was intended to highlight, and precisely what any analysis of narrative must explain. I am hardly the first to suggest as much. For example, Strick and Volbeda (2018) imagined ". . . the classic Romeo and Juliet story without a conflict. Romeo and Juliet would fall in love, their respective clans would immediately support their relationship, and they would get married. End of story. . . . Removing the conflict eliminates the drama" (p. 51).

Visceral Versus Semantic Effects of Stories

There is at least one more matter to be confronted, and that is the possibility that plot may, at times, play a less significant role in engaging the listener than imagined. Recall that the emotional arcs in Figs. 3 and 4 were derived from a story's individual words and therefore make no direct reference to the specific events in the story. They suggest, therefore, that the words employed in a good story could have visceral effects that are potentially independent of plot

---

[12] Insipid television program for young children, a typical "story line" of which is, "The Teletubbies make raspberry noises before watching a little boy called Connor with his grandfather's pigeons" (Episode 350 summary in https://en.wikipedia.org/wiki/List_of_Teletubbies_episodes_and_videos). However, even this program does not fully avoid narrative. For example, in the episode "Windy Day," a hint of dramatic tension is created when the character called Dipsy loses and must recover a prized hat.

developments but could nevertheless modulate the extent to which it is reinforcing to engage with a story.

I have already mentioned how words out of context can evoke emotions. To illustrate how words might be combined to create emotional effects that are separate from a story line, consider the case of abstract poetry, in which words are chosen for aural rather than semantic properties. A case in point is the following excerpt from Edith Sitwell's (1929) poem "Popular Song."

> The red retriever-haired satyr
> Can whine and tease her and flatter,
> But Lily O'Grady,
> Silly and shady,
> In the deep shade is a lazy lady;
> Now Pompey's dead, Homer's read,
> Heliogabalus lost his head,
> And shade is on the brightest wing,
> And dust forbids the bird to sing.
> (reprinted in Sitwell, 1950).

Although it is difficult to say what this passage is about (strictly speaking, it may not be *about* anything), many people feel a sense of poignancy or unease when reading the last line. Such effects have a direct parallel in abstract art, which may lack a definable subject but may nevertheless, through skilled transposition of visual elements, evoke emotion.

These examples suggest that emotion need not be linked only to the semantic properties of verbal behavior. Emotional responses might also be evoked by phonetic and prosodic features of speech[13] and perhaps by the orthographic features of writing (e.g., Grimshaw, Kwasny, Covall, & Johnson, 2003; Kotz et al., 2003). Even nonword and nonspeech vocalizations may evoke listener emotion (Bradley & Lang, 2000b; Schellenberg, Krysciak, & Campbell, 2000). Skinner (1957) anticipated these kinds of effects, acknowledging that some listener emotional responses (for instance, "conditioned reflexes of the Pavlovian variety"; p. 357) are independent of the speaker's putative message ("much of the emotional and imaginal behavior of the listener . . . has little to do with grammar and syntax. An obscene word has its effect regardless of its location or grammar"; p. 344).

The intriguing implication of such effects is that plot might not be as crucial to the success of a story as we like to think. Examples of "plot-optional" stories that have much in common with abstract poetry are easy to suggest. The horror tales of H. P. Lovecraft, for instance, weave a tapestry of disturbing moods but offer limited plot-action. And it might be difficult to apply Hineline's plot-diagramming technique to such popular but famously plot-challenged films as *My Dinner with Andre*, *Lost in Translation*, and *Garden State*. This is not to say that a powerful plot cannot evoke emotion,

---

[13] One of my few vivid memories from an undistinguished English literature undergraduate career is of a professor reading *Beowulf* in Old English in his booming James Earl Jones voice. The fact that I understood barely a word of it did not prevent me from being moved.
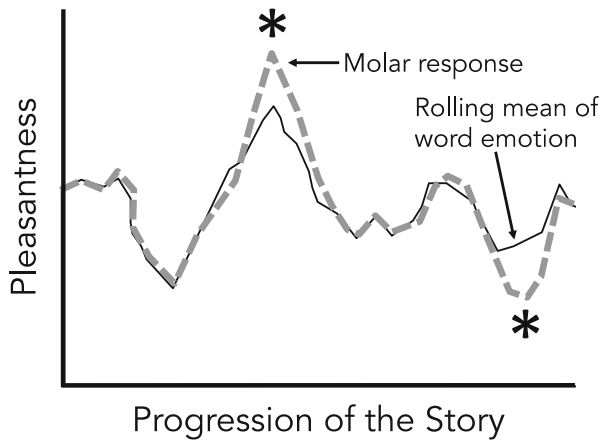
Fig. 5 Comparison between two measures of listener emotion in a hypothetical story. The solid line is a rolling mean of normative emotional responses to the story's component words, and defines an "expected" emotional response. The dashed line is a listener's actual emotional response. Asterisks denote discrepancies between the two measures suggesting that plot points have manipulated motivating operations and potentiated listener emotion. See text for further discussion

only that there may be other ways to evoke emotion that can supplement, or in some instances substitute for, plot.

How, then, in a typical story, might one distinguish between plot-independent emotional effects of words and those aspects of listener emotion that trace to plot? A proper empirical evaluation of any plot "theory" demands this disentangling, and the normative approach of Reagan et al. (2016; see Figs. 3 and 4 in this article) provides the foundation for one strategy of investigation. Figure 5 (solid line) shows, for a hypothetical story, an emotional arc based on an analysis like that of Reagan et al. (2016)—that is, based on emotional responses known to be normatively evoked by the individual words of the story. Call this the story's "expected" emotional arc, or what would occur if all that really matters is listener responses to the story's individual words. The dashed line shows a listener's actual emotional response while digesting the story, perhaps recorded using the "emotion dial" procedure. Note in this example that "expected" and actual emotional responses often are closely correlated, suggesting that, in some places at least, interactive effects of plot elements contribute little to the listener's emotional experience beyond what simple word-emotion can create.[14] But at two points the actual response is magnified compared to what is "expected." The plot developments associated with *these* effects are of particular interest because they imply episodes in which the narrative whole is greater than the sum of its word parts. Such episodes appear to be close to what Hineline (2018) intended to highlight in his analysis of motivating operations and would, I predict, be the most important foci for an empirical study of narrative.

---

[14] This being a convenient what-if discussion, reality is likely to be less straightforward. To mention just two complexities: first, there could be instances in which plot carries the primary burden of listener behavior control, with raw word emotion contributing little. Second, not every individual will mirror normative word emotion, meaning that the "expected" emotional arc is likely to be tricky to determine for individuals, unless raw word emotion data are obtained independently for each of them.

As Hineline (2018) pointed out, in order for behavior science to gain credibility in the public eye, it must examine phenomena that members of society deem important, and an analysis like that of Fig. 5 would not just address behavior scientists' questions about narrative. It would also be of considerable general interest. Mainstream scientists like Reagan et al. (2016; e.g., the present Fig. 4) and various observers of literature (e.g., Vonnegut, 2005) have taken a mostly structural approach to narrative by focusing on the "content" and word selection of particular stories. At best, this approach can indicate how listeners "should" respond to a story (e.g., based on normative word emotion data). The proper focus is on how they actually *do* respond. Although psycholinguists and literature scholars might not instantly resonate to the conceptual framework of behavior science, good measurement of behavior in real time is always in vogue, and a sure way to the hearts of people who are not behavior scientists is to answer questions that they are already asking (Poling, 2010). Outside the confines of behavior science, people know that narrative is important, and they want to understand how it works. If we are willing to do the research—including by astutely selecting good dependent measures—we can show them.

Concluding Observations

Skinner reportedly thought that *Verbal Behavior* (1957) was his most important work. That book presented a conceptual analysis of impressive range and depth but did not address all of the important phenomena to which verbal behavior is integral. Hineline (2018) has done a valuable service by extending a discussion of the potential mechanics of narrative begun earlier by Grant (2005) and Barnes-Holmes and Barnes-Holmes (2002). With these interpretations offering some basic behavioral hypotheses, it is important to begin the arduous task of moving from conceptual to empirical analysis, because the interpretations offered up through conceptual analysis can take a discipline only so far. As Baron, Perone, and Galizio (1991) observed somewhat caustically, nonempirical "interpretations. .. seem to have limited impact, except to generate more interpretations, and perhaps to evoke a sense of self-satisfaction with the apparent scope of the explanatory principle" (p. 102). It is depressing to realize that some 60 years after the publication of *Verbal Behavior* many of the key phenomena it discussed have not yet become the focus of experimental analysis (e.g., Sautter & LeBlanc, 2006) and some of its potentially most important technological implications have yet to be explored (e.g., Critchfield, 2010).

It would be a shame if the same fate befell such a thoughtful analysis as Hineline's (2018), which offers potential insights about the controlling variables of narrative. To set the context for his analysis, Hineline stressed the importance of both a behavioral science and a behavioral technology of narrative, and on this he should be taken at his word. Science and technology advance through data, and I have suggested that it is possible—and desirable—to employ listener emotional responses as a dependent variable in studies of narrative.

This level of analysis may be uncomfortable to those who have been taught to distrust verbal-report measurement and, where empirical research is concerned, variables operating "inside the head." But skeptics should at least consider the potential advantages of tracking listener emotional responses. Doing so may lead to an

unsettling, but fascinating, inversion of the way we normally think about epiphenomena. Remember that Skinner (e.g., 1945, 1953) did not reject emotion as the focus of behavior science on the grounds that it is a fiction. Rather, given the choice between easily observed public behaviors and the emotional responses that are collateral with them, he judged the latter to be more inscrutable, but in the study of narrative the tables may sometimes be turned. Validated means of measuring emotion already exist, but in the context delineated by Hineline (2018) it is not clear whether the same is true for "paying attention," "listening," or "engaging with a story." Some story-related public behaviors such as "spreading favorable reports" fail as dependent measures for other reasons. Where narrative is concerned, therefore, it may be the *nonemotional* responses, co-occurring with a listener's story-related emotions, that are relatively inscrutable. According to Skinner's own logic, *these* would qualify as epiphenomenal.

As Hineline (2018) asserted, storytelling is too prevalent and too influential in human affairs to be left unaddressed by behavior science. Given the importance of the topic, I challenge those opposed to the style of analysis proposed here to do more than just enumerate its presumed shortcomings: if you dislike my focus on emotional responses, come up with a better approach and explain how, to document its superiority, it can be put into practice. That is, consistent with the first rule of storytelling, *show* that you have a better way. Whatever methods are employed, within behavior science the empirical examination of narrative should begin post haste so that a pleasant arc (e.g., "Rags to Riches"?) can one day define the story that historians will tell about how our discipline dealt with this critical topic.

Compliance with Ethical Standards

Conflict of Interest   The author declares declares that he has no conflict of interest.

Ethical Oversight   This work does not draw on primary data collection involving either human or animal participants and thus is not subject to oversight by an ethical review board.

References

Abbott, H. P. (2005). *The Cambridge introduction to narrative* (2nd. ed.). Cambridge: Cambridge University Press.

Aristotle. (2015). The poetics (S. H. Butcher, trans.). London: Macmillan.

Avey, J., Avolio, B. J., & Luthans, F. (2011). Experimentally analyzing the impact of leader positivity on follower positivity and performance. *Leadership Quarterly, 22*, 282–294. https://doi.org/10.1016/j.leaqua.2011.02.004.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97. https://doi.org/10.1901/jaba.1968.1-91.

Barnes-Holmes, Y., & Barnes-Holmes, D. (2002). Naming, story telling, and problem solving: critical elements in the development of language and cognition. *Behavioral Development Bulletin, 1*, 34–38.

Baron, A., Perone, M., & Galizio, M. (1991). Analyzing the reinforcement process at the human level: can application and behavioristic interpretation replace laboratory research? *Behavior Analyst, 14*, 95–105. https://doi.org/10.1007/bf03392557.

Bennett, A., & Royle, N. (2009). *Introduction to literature, criticism, and theory* (4th ed.). New York, NY: Routledge.

Booker, C. (2006). *The seven basic plots: why we tell stories*. New York, NY: Bloomsbury Academic.

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning & Verbal Behavior, 8*, 1–8. https://doi.org/10.1016/s0022-5371(69)80002-2.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Bradley, M. M., & Lang, P. J. (2000a). Measuring emotion: behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York, NY: Oxford University Press.

Bradley, M. M., & Lang, P. J. (2000b). Affective reactions to acoustic stimuli. *Psychophysiology, 37*, 204–215. https://doi.org/10.1111/1469-8986.3720204.

Carr, J. E., Austin, J. L., Britton, L. N., Kellum, K. K., & Bailey, J. S. (1999). As assessment of social validity trends in applied behavior analysis. *Behavioral Interventions, 14*, 223–231. https://doi.org/10.1002/(sici)1099-078x(199910/12)14:4<223::aid-bin37>3.0.co;2-y.

Critchfield, T. S. (2010). Crucial issues in the applied analysis of verbal behavior: a review of *Crucial conversations: tools for talking when the stakes are high*. Analysis of Verbal Behavior, 26, 133–145. doi: https://doi.org/10.1007/bf03393087.

Critchfield, T. S., Tucker, J. A., & Vuchinich, R. E. (1998). Self-report methods. In K. A. Lattal & M. Perone (Eds.), *Handbook of methods for the experimental analysis of human behavior* (pp. 435–470). New York, NY: Plenum Press.

DeGrandpre, R. J., Bickel, W. K., & Higgins, S. T. (1992). Emergent equivalence relations between interoceptive (drugs) and exteroceptive (visual) stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 9–18. https://doi.org/10.1901/jeab.1992.58-9.

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., et al. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences, 112*, 2389–2394. https://doi.org/10.1073/pnas.1411678112.

Estes, W. K., & Skinner, B. F. (1941). Some quantitative properties of anxiety. *Journal of Experimental Psychology, 29*, 390–400. https://doi.org/10.1037/h0062283.

Floh, A., Koller, M., & Zauner, A. (2013). Taking a deeper look at online reviews: asymmetric effect of valence intensity on shopping behavior. *Journal of Marketing Management, 29*, 646–670. https://doi.org/10.1080/0267257x.2013.776620.

Foster-Harris, W. (1959). *The basic patterns of plot*. Norman: University of Oklahoma Press.

Friman, P. C., Hayes, S. C., & Wilson, K. G. (1998). Why behavior analysts should study emotion: the example of anxiety. *Journal of Applied Behavior Analysis, 31*, 137–156. https://doi.org/10.1901/jaba.1998.31-137.

Galton, F. (1908). Suggestions for improving the literary style of scientific memoirs. *Transactions of the Royal Society of Literature in the United Kingdom (Second Series), 27*, 113–130.

Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality & Social Psychology Bulletin, 30*, 585–593. https://doi.org/10.1177/0146167203262856.

Gonzalez, R. (2014). The shapes of stories, according to Kurt Vonnegut. IO9. Retrieved from https://io9.gizmodo.com/the-universal-shapes-of-stories-according-to-kurt-vonn-1526559996

Grant, L. K. (2005). The secrets of Scheherazade: toward a functional analysis of imaginative literature. *Analysis of Verbal Behavior, 21*, 181–190. https://doi.org/10.1007/bf03393020.

Grimshaw, G. M., Kwasny, K. M., Covall, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia, 41*, 1008–1019. https://doi.org/10.1016/s0028-3932(02)00315-9.

Hineline, P. N. (2018). Narrative: what it is and how it works. *Perspectives on Behavior Science*. https://doi.org/10.1007/s40614-018-0137-x.

Izenson, B. (2016). Dial testing called back into duty to deliver real-time returns on the Presidential debates. Dialsmith. Retrieved from http://www.dialsmith.com/blog/dial-testing-called-back-duty-to-deliver-real-time-returns-for-presidential-debates/).

Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of behavioral research*. Upper Saddle River, NJ: Lawrence Erlbaum Associates.

Kotz, S. J., Meer, M., Altr, K., Besson, M., von Cramon, D. Y., & Friederici, A. A. (2003). On the lateralization of emotion prosody: an event-related functional MR investigation. *Brain & Language, 86*, 366–376. https://doi.org/10.1016/s0093-934x(02)00532-1.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological Psychiatry, 44*, 1248–1263. https://doi.org/10.1016/s0006-3223(98)00275-3.

Lubbock, P. (1921). *The craft of fiction*. New York, NY: Charles Scribner's Sons.

Norman, S. M., Avolio, B. J., & Luthans, F. (2010). The impact of positivity and transparency on trust in leaders and their perceived effectiveness. *Leadership Quarterly, 21*, 350–364. https://doi.org/10.1016/j.leaqua.2010.03.002.

Petty, R. E., Schumann, D. W., Richman, S. S., & Strathman, A. J. (1993). Positive mood and persuasion: different roles for affect under high- and low-elaboration conditions. *Journal of Personality & Social Psychology, 64*, 5–20. https://doi.org/10.1037/0022-3514.64.1.5.

Poling, A. (2010). Looking to the future: Will behavior analysis survive and prosper? *Behavior Analyst, 33*, 7–17.

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. EPJ Data Science, 5, Article 31. doi: https://doi.org/10.1140/epjds/s13688-016-0093-1.

Rogers, C. R., & Skinner, B. F. (1956). Some issues concerning the control of human behavior: a symposium. *Science, 124*, 1057–1066. https://doi.org/10.1126/science.124.3231.1057.

Rowling, J. K. (2009). *Harry potter and the deathly hallows*. New York, NY: Arthur A. Levine.

Sautter, R. A., & LeBlanc, L. A. (2006). Empirical applications of Skinner's analysis of verbal behavior with humans. *Analysis of Verbal Behavior, 22*, 35–48. https://doi.org/10.1007/bf03393025.

Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving emotion in melody: interactive effects of pitch and rhythm. *Music Perception, 18*, 155–171. https://doi.org/10.2307/40285907.

Sidman, M. (1960). *Tactics of scientific research*. New York, NY: Basic Books.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127–146.

Sitwell, E. (1929). *Popular song*. London: Faber & Faber, Ltd..

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review, 52*, 270–277. https://doi.org/10.1037/h0062535.

Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Free Press.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Stadthagen-Gonzalez, H., Imbault, C., Perez Sanchez, M. A., & Brysbaert, M. (2016). *Norms of valence and arousal for 14,031 Spanish words. Behavior Research Methods*. Online prepublication. https://doi.org/10.3758/2Fs13428-015-0700-2.

Strick, M., & Volbeda, M. (2018). When the valence of unconditioned stimuli evolves over time: evaluative conditioning with good-ending and bad-ending stories. *Journal of Experimental Social Psychology, 74*, 50–55. https://doi.org/10.1016/j.jesp.2017.08.006.

Tobias, R. B. (1993). *20 master plots and how to build them*. Blue Ash, OH: Writer's Digest Books.

Vonnegut, K. (2005). At the blackboard. Lapham's Quarterly. Retrieved from https://www.laphamsquarterly.org/arts-letters/blackboard

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*, 1–17. https://doi.org/10.3758/s13428-012-0314-x.

Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214. https://doi.org/10.1901/jaba.1978.11-203.