

The Importance of Morally Satisfying Endings: Cognitive Influences on Storytelling in Gillian Flynn's *Gone Girl*

Sarah G. P. Binau (sarah.binau@pomona.edu)

Department of Linguistics and Cognitive Science, Pomona College
333 N. College Way, Claremont, CA 91711 USA

Robin J. Melnick (robin.melnick@pomona.edu)

Department of Linguistics and Cognitive Science, Pomona College
333 N. College Way, Claremont, CA 91711 USA

Jack I. Abecassis (jack.abecassis@pomona.edu)

Department of Romance Languages, Pomona College
333 N. College Way, Claremont, CA 91711 USA

Abstract

Peak End Rule (Kahneman, 1993; 2011) suggests that the average of the peak and end moments of an event disproportionately affect memory and thus perception of the experience. We investigate PER's application to the experience of reading fiction. Gillian Flynn's *Gone Girl* (2012) is an ideal case study because it is commercially popular but, unlike most popular novels, has a distinctly amoral ending. We hypothesize that humans expect moral payoffs at the end of narrative fiction, and that when these expectations are not met (i.e., pain at the end of the experience), as in the case of *Gone Girl*, readers' perceptions of the story will be influenced by this pain and manifest as disappointment and dislike. We reference existing models in evolutionary psychology, which seek to explain human altruism, and models in cognitive science, which seek to explain patterns in memory and assessment. To quantify disappointment and dislike, we conduct a programmatic corpus linguistic analysis of 40,000 web-scraped Amazon product reviews of *Gone Girl*, comparing them to reviews of eight other similarly popular novels from the same year. Results show that reader sentiments about *Gone Girl*, both the overall review ratings and analysis on a sentence-by-sentence basis, are more positive than for the comparison novels. When only reviews mentioning "end" are analyzed, however, the effect reverses, with a similar finding at the more granular level of sentences mentioning "end." These findings support our hypothesis that moral endings, or lack thereof, significantly shape reader perceptions of a novel.

Keywords: peak end rule; narrative endings; sentiment analysis; corpus linguistics; web scraping; Amazon product reviews; morality in narrative; evolutionarily stable systems; social cooperation; *Gone Girl*

Introduction

The principle of Peak End Rule (PER) suggests that a memory of an experience is influenced disproportionately by two key moments: the most intense moment of pain in the episode and the level of pain felt at the end of the episode (Kahneman et al., 1993). Multiple experiments have supported the notion of PER, utilizing a variety of methodologies (Fredrickson & Kahneman, 1993; Kahneman, 2011; Redelmeir & Kahneman, 1996). While Kahneman (2011) studied how subjects evaluate fictional stories of people's lives (ending happily or sadly) in a Peak End Rule

framework, it remains to be explored how PER applies to readers' perceptions of literary fiction. In the present study, we explore such perceptions as expressed through reader book reviews. Here, the event to which PER applies is the reading of a novel. Since we assume that such review assessments are highly correlated to perceptions formed during the reading event, perception and memory here are then intertwined.

Literary critics working within an evolutionary framework believe fiction to be inherently moral (Gardner, 1978; Gottschall, 2012). Regardless of how much readers delight in highly problematic narratives, they nevertheless expect a moral payoff by way of closure (Carroll, 2011; Flesch, 2009; Gardner, 1978). Moreover, it has been shown that authors and filmmakers are acutely aware of this expectation and use cognitive biases in the construction of narratives to manipulate the effect on readers (Smith, 2015).

This expectation for moral endings may be rooted in the evolutionary advantages of altruism. Altruistic action is necessary in group systems because it ultimately promulgates group stability (Flesch, 2009). In Evolutionarily Stable Systems, people are cooperators, defectors, or punishers (Flesch, 2009). Cooperators constantly track the behavior of others in their group using gossip, honest signaling, and other social tools to ensure that no one is defecting (Dunbar, 2004; Flesch, 2009; Harari, 2015; Zahavi & Zahavi, 1997). If cooperators catch defectors, they seek to punish the latter by exposing their free-riding behavior to the rest of the group. This same human propensity for altruistic punishment has been demonstrated through the game theory constructs of the Ultimatum Game and the Prisoner's Dilemma (Fehr & Fischbacher, 2003; Güth et al., 1982; Tucker, 1983).

William Flesch argues that the moment the author allows moral characters to expose and punish immoral characters is the "pleasure of fiction" (Flesch, 2009). In fact, it is not just the authors' ability, but their social responsibility to provide this payoff (Carroll, 2011). Reader expectations for this payoff are also influenced by genre: when viewers watch more fictional television, they develop stronger Just World Beliefs, as opposed to when they watch news/infotainment

television (Appel, 2009). This could be because fictional narratives “tend to stimulate moral evaluation,” whereas nonfictional television is saturated with immoral behaviors not necessarily followed by moral consequences (Appel, 2009). In this way, genre affects readers’ expectations.

A case study: *Gone Girl*

Using the novel *Gone Girl* by Gillian Flynn (2012), we present an emblematic case study for moral endings in fictional narratives. *Gone Girl* is a useful case in that it is commercially popular, topping *New York Times* (NYT) bestseller lists and garnering over 40,000 product reviews on Amazon, but its ending is highly disturbing because of its lack of morality (Amazon Customer Reviews, 2016; *New York Times* Bestsellers, 2012, 2013).

For readers to engage with a book, they inherently need to trust the narrator (Carroll, 2011). Flynn expertly manipulates this trust to deliver a spectacularly amoral ending, thus experienced by the readers as particularly painful. With two unreliable narrators, Flynn’s readers perceive their highly manipulated empathy for the two protagonists, Nick and Amy, as wasted. In exchange for this deception, readers expect Flynn to craft an even stronger, satisfying ending.

However, the novel’s ending is both abrupt and lacking in a moral payoff. The book has no altruistic punishers. The defector, Amy, is not punished sufficiently; she succeeds in all of her free-riding efforts. As one reader comments: “I was disappointed in the ending. I was hoping for either justice on Nick’s side or the demise of Amy. Ending just wasn’t what I had hoped it would be.” Given the principle of Peak End Rule, we would expect readers’ perceptions and global assessments of the novel to be most influenced by the painful and unsatisfying ending (Kahneman, 2011).

To assess this and other related claims about moral closure in novels, this paper linguistically analyzes Amazon product reviews of *Gone Girl* (Amazon Customer Reviews, 2016). This methodology is derived from an increasing usage of online corpora to assess opinions of a product or work (Allington, 2016; Boot, 2013).

Hypotheses

Readers expect moral endings out of *New York Times* bestselling narrative fiction, and when these expectations are not met, as is the case in *Gone Girl*, readers will feel dissatisfied and disappointed because their unspoken contract with the author has been broken. Specifically, we may expect to observe the following:

(1) Readers of *Gone Girl* will more frequently discuss the ending, as compared to other best-selling literary novels.

(2) While *Gone Girl* is similarly popular (i.e., best-selling), with comparable overall Amazon review numerical ratings (scale: 1-5 “stars”), reviews discussing the “end” will garner significantly lower ratings for *Gone Girl*. Discussing “end” is not expected to similarly affect the comparison corpus.

(3) At the sentential level within *Gone Girl* reviews, mentioning “end” will be associated with significantly more negative surrounding sentiment, as judged by a machine-learning classifier. No similar effect is expected for the comparison corpus.

(4) Descriptively, adjectival collocates of “end” are expected to be substantially more negative within the *Gone Girl* corpus.

Methods

Two corpora were created and analyzed using a combination of web-scraping tools, custom scripting software, part-of-speech tagging, deep-learning sentiment classification, and the R statistics platform (version 3.5.1, R Core Team, 2018). Throughout this work, we consider references to the ending of a novel to be any of “end,” “ends,” “ended,” or “ending,” referring to this set collectively as simply “end.”

Materials

Star-rated consumer reviews of novels were extracted from the Amazon website and organized into two corpora: one for the novel *Gone Girl* (Flynn, 2012) and one for a group of comparison works.

***Gone Girl* corpus** A total of 39,436 product reviews of the novel *Gone Girl* by Gillian Flynn were extracted in December 2016 using ParseHub web-scraping software. Data collected were: title of review, content of review, date published, and star rating. After a pilot study of the full corpus, final analysis was limited to the first 2,000 reviews chronologically following the book’s release, all within 2012, in order to avoid any influence of the movie *Gone Girl* (Fincher, 2014), announced in 2013 and released in 2014.

Comparison corpus The comparison group of novels comprises all works that, like *Gone Girl*, appeared on the NYT bestselling fiction list for two or more consecutive weeks in 2012 or 2013. Our collected corpus contains product reviews for: *11/22/63* by Stephen King, *Fifty Shades of Grey* by E.L. James, *Reflected in You* by Sylvia Day, *The Racketeer* by John Grisham, *A Memory of Light* by Robert Jordan and Brandon Sanderson, *Until the End of Time* by Danielle Steel, *Inferno* by Dan Brown, and *The Cuckoo’s Calling* by Robert Galbraith. Amazon reviews were extracted, collecting title of book, plus as before: title of review, content of review, date published, star rating.

Here, the number of reviews extracted for each novel was limited to either the first 2,000 chronologically or all reviews appearing within the first two years of release, whichever was less. As with *Gone Girl*, this was done to limit any potential influence from television or cinematic releases based on the books. Within these parameters, a total of 14,460 reviews were extracted, *11/22/63* and *Until the End of Time* via ParseHub, with the remainder collected using the webscraper.io utility (Balodis, 2018).

All tokens of the phrase “until the end of time,” without regard to case, were excluded from analysis, as this is not only a somewhat fixed multi-word expression (MWE) in English, but here it is also specifically the title of one of our comparison novels. We assume the use of “end” in this expression does not actually refer to the ending of the narrative. A second fixed MWE, “loose ends,” was similarly excluded.

Results

Use of *end*

Per hypothesis 1, reviewers of *Gone Girl* discuss the ending more. In the comparison corpus, 25.7% of the reviews mention “end,” 3.7% of the review titles, and “end” comprises 0.21% of all words. The corresponding figures for *Gone Girl* are 52.4%, 8.4%, and 0.63%, more than double in each case (Figure 1).

Differences by “star” rating

Hypothesis 2 predicts that mentioning “end” will lower Amazon review numerical ratings for *Gone Girl* more so than for the comparison set. Figure 2 appears to bear this out.

To explore this further, we fit a linear mixed-effects regression model to predict review star rating from fixed effects for source corpus, length of review (in words, log-reduced to limit outlier effects), and whether or not each review included mention of “end” in its title or body text, with a random effect for individual book title.¹ We trialed all

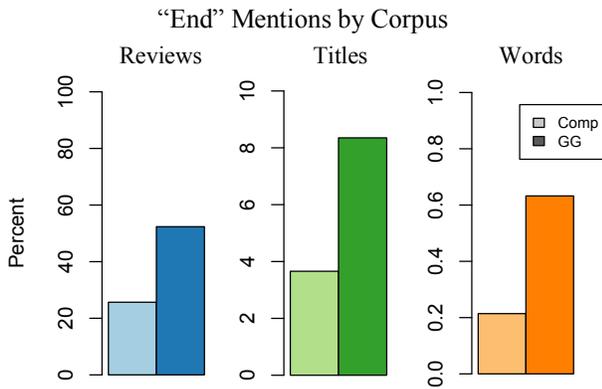


Figure 1: Percentages of “end” mentions for comparison corpus (“Comp”, lighter bars) vs. *Gone Girl* corpus (“GG”, darker bars).

¹ We do not also include a random effect for “participant” (i.e., reviewer) since, with rare exception, there is just one review per reviewer available within our pair of corpora—that is, there is no clustering to model.

² The model exhibits negligible collinearity, with condition number $\kappa = 1.41$ (where Cheney and Kincaid 2007 suggests 10+ is problematic) and largest variance inflation factor (VIF) for any single predictor = 1.31 (10+ again being high, per Hair et al. 1998).

³ Inclusion of length in this analysis was motivated by an anonymous commenter’s concern that lower ratings for *Gone Girl* reviews mentioning “end” might arise if these reviews tended to be

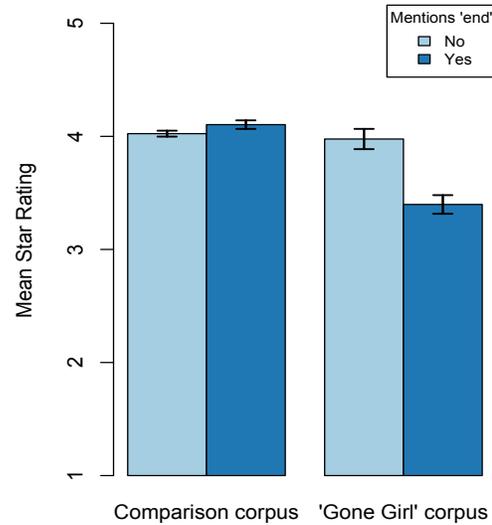


Figure 2: Effects of mentioning “end” on overall reviewer rating. (Comparison corpus $N=14.5K$, *Gone Girl* $N=2K$.)

possible two-way interactions among the fixed effects, as well as a maximal random effects structure (Barr et al., 2013). Stepwise model optimization preserves the random intercept for book title, the associated random slope for review length, and interactions of source corpus with both review length and “end” mentions (Table 1).²

As a baseline, there was no significant difference in overall star rating among these popular works when “end” was not mentioned (main effect of source corpus, $\hat{\beta} = 0.019$, $t = 0.078$, $p = 0.9403$), and longer reviews correlated with somewhat lower ratings overall ($\hat{\beta} = -0.283$, $t = -7.182$, $p = 0.0002$).³ When “end” is mentioned, mean ratings

Table 1: Fixed effects and interactions from linear regression, with baseline y -intercept of about 4 out of 5 stars.

	Coef. $\hat{\beta}$	SE($\hat{\beta}$)	t	$Pr(> t)$
Intercept	3.964	0.081	49.181	< 0.0001
Corpus: <i>Gone Girl</i>	0.019	0.242	0.078	0.9403
Review length	-0.283	0.039	-7.182	0.0002
Mentions “end”	0.260	0.026	10.145	< 0.0001
<i>Gone Girl</i> \times length	0.353	0.118	3.006	0.0197
<i>Gone Girl</i> \times “end”	-0.870	0.064	-13.606	< 0.0001

longer, where longer reviews might in turn correlate with lower ratings. As reported above, the main effect of review length does turn out to be correlated with lower scores overall, but we find that greater length specifically among the *Gone Girl* reviews actually positively influences star rating (interaction of source corpus and length, $\hat{\beta} = 0.353$, $t = 3.006$, $p = 0.0197$). Further, *Gone Girl* reviews overall are not significantly longer than the comparison group ($\bar{x} = 101.5$ words *GG* vs. 100.6 *comp*, $t = 0.322$, $p = 0.7473$), and *Gone Girl* reviews mentioning “end” are actually significantly shorter on average than reviews mentioning “end” within the comparison corpus ($\bar{x} = 123.0$ vs. 160.4, $t = -7.027$, $p < 0.0001$).

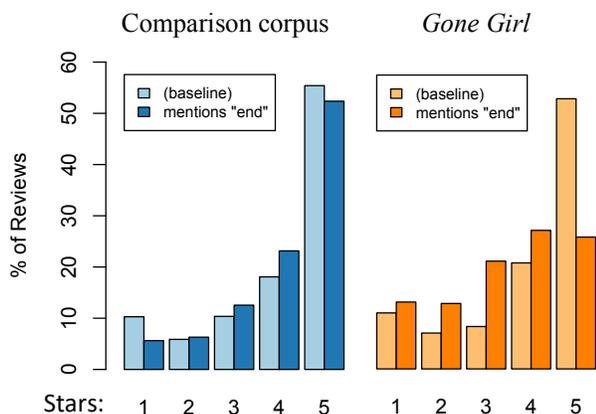


Figure 3: Star ratings by source, with and without “end.”

actually increase overall ($\hat{\beta} = 0.260$, $t = 10.145$, $p < 0.0001$), but looking at the interaction terms, among *Gone Girl* reviews that mention “end,” scores significantly decline ($\hat{\beta} = -0.870$, $t = -13.606$, $p < 0.0001$) as predicted.

Figure 3 further illustrates the distribution of star ratings with and without “end” mentions.

Sentiment analysis

The analyses detailed above offer a view of how mentioning “end” is reflected in an extrinsic measure of overall reviewer satisfaction—the star ratings. But a question remains with regard to what reviewers express more directly about the ending of the story.

To explore such an intrinsic measure, we began by splitting review text at punctuation points, a proxy for clause boundaries, yielding 136K such segments, including 12,277 for *Gone Girl*, of which 1,066 mention “end.” We applied deep-learning sentiment annotation via the Stanford CoreNLP toolkit (Socher et al., 2013; Manning et al. 2014) to rate the emotional content of each segment of text (“clause”) as “Very negative,” “Negative,” “Neutral,” “Positive,” or “Very positive.”

Recasting these ratings as a continuous scale (-2 to +2), we once again fit a linear mixed-effects regression model, now predicting sentiment score from fixed effects for source corpus, clause length, and whether or not each such individual segment of text included mention of “end.” To these, we added random effects, now both for book title and for individual review.⁴ As before, we also model all possible two-way interactions of the fixed effects and the maximal random-effects structure supported by the data. Stepwise optimization this time preserved all main effects and their interactions (Table 2), as well as both of the random intercepts and a random slope for the separate effect of “end”-mention within each given review.⁵

⁴ Whereas our previous analysis—star ratings applied to the full reviews—did not include a per-review (i.e., “participant”) effect.

Table 2: Fixed effects and interactions from regression, with baseline y-intercept = 0.02 (i.e., overall “Neutral” sentiment).

	Coef. $\hat{\beta}$	SE($\hat{\beta}$)	t	$Pr(> t)$
Intercept	0.020	0.038	0.515	0.6223
Corpus: <i>Gone Girl</i>	0.068	0.114	0.593	0.5720
Clausal length	-0.181	0.003	-61.121	< 0.0001
Mentions “end”	0.052	0.021	2.462	0.0139
<i>Gone Girl</i> × “end”	-0.351	0.041	-8.575	< 0.0001
<i>Gone Girl</i> × Length	0.025	0.010	2.406	0.0161
“End” × Length	-0.048	0.022	-2.149	0.0317

At this clausal level, there was again no significant main effect of source corpus, meaning segments of *Gone Girl* reviews overall were neither more positive nor more negative than those in the comparison corpus ($\hat{\beta} = 0.068$, $t = 0.593$, $p = 0.5720$). We also find that longer clauses are more likely to bear negative sentiment (main effect of length, $\hat{\beta} = -0.181$, $t = -61.121$, $p < 0.0001$), much as our previous analysis found that longer complete reviews received lower star ratings. This effect was, however, weakened within *Gone Girl* review prose (interaction term, $\hat{\beta} = 0.025$, $t = 2.406$, $p = 0.0161$). Among all clauses mentioning “end,” mean sentiment increases overall ($\hat{\beta} = 0.052$, $t = 2.462$, $p = 0.0139$), much as did full-review star ratings. Critically, though, *Gone Girl* “end”-mentions are significantly more negative (interaction $\hat{\beta} = -0.351$, $t = -8.575$, $p < 0.0001$), just as we saw with lower star ratings for full reviews, again as predicted by Hypothesis 3.

Most common descriptive terminology

Finally, we examined adjectival collocates of “end” to see how reviewers specifically describe the respective endings. Here, we began by applying part-of-speech (POS) labels to all text, once again using Stanford CoreNLP (Toutanova et al., 2003; Manning et al., 2014), then analyzed adjectives appearing within a three-word window of “end,” left or right, without crossing clausal or sentential boundaries. Negated contexts (e.g., “not good,” “never disappointing”) were excluded from this portion of the analysis (Pang et al., 2002).

The most frequent adjective used to describe “end” in the *Gone Girl* corpus was “disappointing,” which in combination with the related form “disappointed” was almost twice as frequent as the next most used word to describe the ending of the story, “worst” (Table 3).

Examples from the *Gone Girl* reviews of the use of “disappointing” or “disappointed” in relation to “end” include: “the ending was disappointing to the point of making me wish I had not spent my time reading this,” “it was a good book, but the ending was a disappointment,” and “the big disappointment was the ending...this book ended horribly.”

⁵ As with our earlier model, we find negligible collinearity, with $\kappa = 1.15$ and largest VIF for any single predictor = 1.77.

Table 3: Most common adjectives used in descriptions of “end” in the *Gone Girl* corpus. Coding for sentiment as rated by the Stanford CoreNLP toolkit (Socher et al., 2013; Manning et al. 2014): red/- = negative; green/+ = positive; gray/~ = neutral.

Adjective	Frequency (%)
- disappointing	5.08
- worst	3.81
+ good	3.60
- horrible	2.97
- unsatisfying	2.97
~ little	2.75
- awful	2.54
+ happy	2.54
- terrible	2.33
- bad	2.12
- disappointed	1.91

The adjective most frequently used to describe the ending in the comparison corpus was “great” (Table 4). Excerpts from reviews including use of “great” to describe “end” in the comparison corpus include: “great story...with a literally killer climax and great ending,” “what a great ending to a terrific trilogy,” “the ending was great and totally unexpected,” and “awesome book, great ending to an epic saga.”

While these respective sets of most frequent adjectives appear largely disjoint, we find a few terms in common among those frequently applied in both corpora, e.g., “good” and “little.” To then further explore which terms most distinctively apply to “end” in one corpus as compared to the other, we examined relative frequency across the two data sets. We log-reduced values and Z-score normalized for comparison, then found the ratio of “end”-description frequency for *Gone Girl* over the comparison set. The largest ratios (Table 5) represent terms most frequently applied to “end” for *Gone Girl* vs. the comparison group.

Table 4: Most common adjectives used in descriptions of the endings of comparison novels. (Sentiment coding as in Table 3.)

Adjective	Frequency (%)
+ great	12.36
+ good	5.53
+ happy	3.76
+ perfect	2.69
+ wonderful	2.23
~ little	2.07
+ amazing	2.00
+ satisfying	2.00
+ fantastic	1.69
~ first	1.69
+ excellent	1.53

Table 5: Adjectives most distinctively applied to “end” in *Gone Girl* vs. comparison novels, as measured by relative frequency across the two. (Sentiment coding as in Table 3.)

Adjective	GG/Comp ratio
- awful	24.54
- worst	18.40
- anticlimactic	15.34
- flat	12.27
- bizarre	9.20
- disappointed	9.20
- unsatisfying	8.18
- ridiculous	7.67
- atrocious	6.13
- terrible	6.13
- horrible	6.13

Gone Girl corpus examples of “awful”—24.54 times more likely to be used to describe “end” for *Gone Girl* than for the comparison set—include: “slow pace and awful ending—don’t waste your money,” “great book, awful ending,” “it could’ve been a perfect novel, but that God awful ending!” and “awful ending...makes you feel dirty for wasting so much time on a sick, twisted book.”

Finally, we further visualized this notion of relative frequency ratio, focusing on the 20 adjectives with greatest average “end”-collocation frequency across the two corpora. Figure 4 graphs their respective (log-reduced, normalized) frequencies for the two data sets, with sentiment labeling once again as above in Table 3. Terms below the dividing line were more frequently “end” collocates for *Gone Girl*, none expressing positive sentiment. Those above the diagonal, none of which express negative sentiment, were more frequently applied to “end” for comparison works.

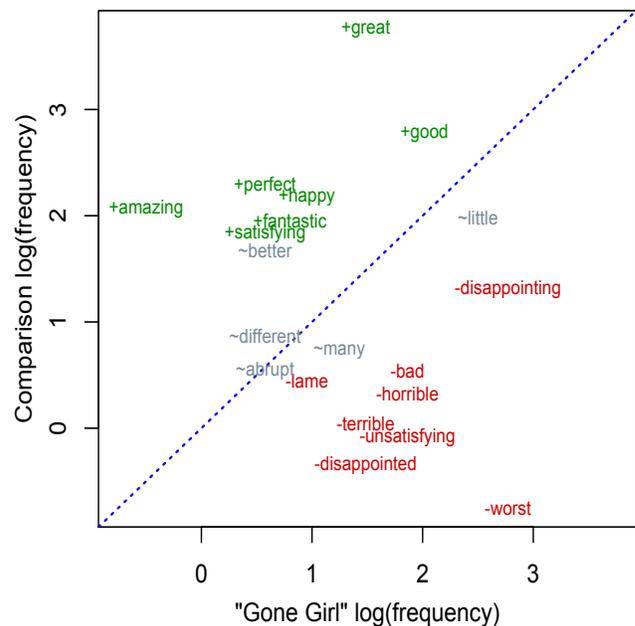


Figure 4: Relative frequency of collocates across corpora.

Discussion

Moral payoffs are closely related to the pleasure of fiction. Readers make an unspoken contract with the author before reading a book: I will read your book, expecting that the morally problematic content will be cleaned up by the end. When this baseline assumption is violated, as is the case in *Gone Girl*, the ending becomes a salient discussion topic. Our findings confirm this: mentions of “end” were twice as common in *Gone Girl* reviews than comparison novel reviews.

And readers do not merely discuss the ending: they vividly express their disappointment. In fact, “disappointed” was used almost 10 times more often in relation to the ending of *Gone Girl* than for the comparison novels. This adjective in particular is revealing in the context of the reader/author contract, because it signifies readers’ expectations and their lack of fulfillment.

In contrast, the adjectives most frequently used to describe “end” in our comparison group of novels were generally positive, suggesting satisfaction with the ending and, therefore, fulfillment of the unspoken contract with the author.

Gone Girl is an important case study in the expectation for moral payoffs in novels because it lacks a moral ending, yet it is commercially popular. In fact, in terms of overall review scores, *Gone Girl* is just as highly rated as comparison best-selling novels. What stands out, however, is that when reviews are separated into those that mention “end” and those that do not, *Gone Girl* reviews significantly differ from comparison reviews. Crucially, we found that while including a mention of “end” in the comparison corpus does not significantly affect star rating, mentioning “end” in the *Gone Girl* corpus significantly lowers ratings. This suggests that *Gone Girl*’s amoral ending is a salient cause for overall dissatisfaction with the novel.

While this conclusion is made on the basis of reviewers’ overall star ratings, an extrinsic measure, we found similar evidence in programmatic sentiment analysis of review prose. Here, the effect of mentioning “end” was even more pronounced: *Gone Girl* commentary about matters other than “end” was significantly more positive than in reviews of comparison works, while discussions of “end” were significantly more negative.

The fact that discussions of “end” were extremely negative in the *Gone Girl* corpus suggests that reviewers act as altruistic punishers in the framework of Evolutionary Stable Systems. Because Flynn controls the narration that explains the world of *Gone Girl*, readers expect her to act as the primary altruistic punisher in the novel’s social ecosystem. If there is a defector within the novel, it is Flynn’s job to guide the narration such that the defector is exposed and consequently punished. In the case of *Gone Girl*, there is more than one defector, and Flynn fails to craft the narration to punish any of them appropriately. Readers experience and react to Flynn’s lack of punishment and become second-order altruistic punishers themselves; they go to Amazon and write product reviews for the novel, explaining their opinions of the

ending and exposing Flynn herself as a defector for her failure to include a moral ending.

Therefore, we see that the evolutionary advantages of altruistic behavior not only guide immediate social groups, but literary communities as well. The principle of Peak End Rule further compounds this: when psychologically amoral, painful moments occur at the end of experiences or narratives, the impact is amplified. This heightening of moral discomfort is what drives readers to write over 40,000 product reviews for a novel, over half of which include a discussion of the “end,” with “disappointing” as a key descriptor.

Future Directions

Given the wealth of data available via Amazon, further investigation using more novels and more reviews is an important next step. We would also like to consider reviews from other Internet sources, such as GoodReads.

Another possibility would be to examine responses to the movie *Gone Girl*, looking at how these compare with reactions to the novel. Although Flynn wrote the screenplay for the movie, its narration style inherently changes with the shift in medium, meaning that reactions to the ending could shift as well.

Conclusion

The Peak End Rule principle suggests that endings of experiences or narratives significantly affect a person’s overall memory, or perception, of that event. Though first explored through cognitive research on human memory, the principle also applies to readers’ perceptions of literary novels, and in particular, their endings. That is, moral endings, or lack thereof, as in the case of *Gone Girl*, have a strong effect on readers’ perceptions of novels.

In light of the human propensity for cooperation, *Gone Girl*’s lack of a moral ending dramatically affects reader response, in comparison with other popular contemporary novels. This investigation demonstrates the profound effect endings have in shaping conception of stories, as well as our expectation for a moral payoff in literary novels—even disturbing ones.

Acknowledgements

We are grateful for the helpful comments from anonymous reviewers and from the attendees of a presentation of this work at Pomona College. Any errors remain our own. Statistical calculations and graphics were developed with R version 3.5.1 (R Core Team, 2018).

References

- Allington, D. (2016). ‘Power to the reader’ or ‘degradation of literary taste’? Professional critics and Amazon customers as reviewers of *The Inheritance of Loss*. *Language and Literature*. 25(3): 254–278.
- Amazon Customer Reviews. (2016). *Gone Girl* by Gillian Flynn. Retrieved from <http://a.co/d/4MPuLjL>.

- Appel, M. (2008). Fictional narratives cultivate just-world beliefs. *Journal of Communication*, 58: 62-83.
- Balodis, M. (2018). Web Scraper [Computer software]. Retrieved from <https://chrome.google.com/webstore>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Boot, P. (2013). The desirability of a corpus of online book responses. In *Proceedings of the Workshop on Computational Linguistics for Literature* (pp. 32-40). Association for Computational Linguistics.
- Carroll, N. (2011). Narration. In P. Livingston & C. Plantinga (Eds), *The Routledge Companion to Philosophy and Film*. London: Routledge.
- Cheney, E. W., & Kincaid, D. R. (2012). *Numerical Mathematics and Computing*. Cengage Learning.
- Coplan, A. (2011). Empathy and character engagement. In P. Livingston & C. Plantinga (Eds), *The Routledge Companion to Philosophy and Film*. London: Routledge.
- Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, 8(2), 100-110.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature International Journal of Science*, 425, 85-791.
- Fincher, D. (Producer/Director). (2014). *Gone Girl* [Motion picture]. United States: Twentieth Century Fox.
- Flesch, W. (2009). *Comeuppance*. Harvard University Press.
- Flynn, G. (2012). *Gone Girl: A Novel*. NYC: Crown.
- Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, 65(1), 45-55.
- Gardner, J. (1978). *On Moral Fiction*. NY: Basic Books.
- Gottschall, J. (2012). Morality. *The Storytelling Animal: How Stories Make Us Human*. Boston: Houghton Mifflin Harcourt, 117-138.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3, 367-388.
- Hair, J., Andreson, R., Tatham, R., & Black, W. (1998). *Multivariate Data Analysis* (5th ed.). Prentice-Hall, Inc.
- Harari, Y. (2015). The tree of knowledge. In *Sapiens: A Brief History of Humankind*. Purcell, J., & Watzman, H., Trans. New York: Harper, 20-39.
- Kahneman, D., Fredrickson, B., Schreiber, C., & Redelmeier, D. (1993). When more pain is preferred to less: Adding a Better End. *Psychological Science*, 4(6): 401-05.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.
- Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.
- New York Times Bestsellers. (2013). *Fiction*. Retrieved from <https://www.nytimes.com/books/best-sellers/>.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing* (pp. 79-86). Association for Computational Linguistics.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Redelmeier, D., & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain*. 66 (1): 3-8.
- Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1631-1642). Association for Computational Linguistics.
- Smith, J. (2015). Filmmakers as folk psychologists: How filmmakers exploit cognitive biases as an aspect of cinematic narration, characterization, and spectatorship. In *The Oxford Handbook of Cognitive Literary Studies*. Oxford University Press.
- Toutanova, K., Klein, D., Manning, C., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of HLT-NAACL 2003*, pp. 252-259.
- Tucker, A. (1983). The mathematics of Tucker: A sampler. *Mathematical Association of America*. 14 (3): 228.
- Zahavi, A., & Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. New York: Oxford University Press.