





The Authorship of Stephen King's Books Written Under the Pseudonym "Richard Bachman": A Stylometric Analysis

Dorothy Modrall Sperling¹ 

Vincent Neyt² 

Mike Kestemont² 

1. Karel De Grote University College , Antwerp, Belgium.

2. Department of Literature, University of Antwerp , Antwerp, Belgium.

Citation

Dorothy Modrall Sperling, Mike Kestemont, and Vincent Neyt (2023). "The Authorship of Stephen King's Books Written Under the Pseudonym "Richard Bachman". A Stylometric Analysis". In: *Journal of Computational Literary Studies* 2 (1). [10.48694/jcls.3594](https://doi.org/10.48694/jcls.3594)

Date published 2024-03-02

Date accepted 2023-12-20

Date received 2023-01-31

Keywords

Stephen King, stylometry, pop culture, authorship attribution, contemporary English-language fiction

License

CC BY 4.0 

Reviewers

Erik Ketzan, Jacques Savoy

Note

This paper has passed through the conference track of JCLS. In addition to being peer reviewed, it was presented and discussed at the 2nd Annual Conference of Computational Literary Studies at Würzburg University in June 2023.

Abstract. Between 1977 and 1984, Stephen King published five novels under the pseudonym "Richard Bachman". Reviewers noted similarities between King's and Bachman's writing styles when *Thinner* (1984) was published, ultimately leading to King's unmasking. We investigate, using the Juola protocol, whether computational techniques can correctly identify King as the author of the Bachman books out of a selection of contemporary candidate authors – Dean Koontz, Peter Straub, and Thomas Harris. We also perform a post-hoc analysis of the use of pop-culture references and brand names in Bachman, King, Koontz, Straub, and Harris novels, based on comments in reviews of Bachman and King novels. The references extracted from the Bachman books occurred significantly more often in King's texts than in the others', showing that attentive readers could have "heard King's voice" in the Bachman books through what a reviewer denigratingly called King's "compulsion to list brand-name products and his affinity for pop-cult teenage junk". These results contribute to the vexed issue of explainability, which is a recurrent challenge in author identification for literary texts.

1. Introduction

In February 1980, Stephen King published a lengthy essay called "On Becoming a Brand Name" in *Adelina*, a men's magazine. The essay records King's reaction to being referred to as a "Brand Name Author" in the modern American horror genre. He did not wish to oppose himself to being given that label; he was proud of his accomplishments at that point as a regular in the *New York Times* bestseller list and receiver of advances of well over one million dollars for a new novel. Instead, he wrote the essay to tell the story of how he had arrived at being labelled as such, emphasizing especially that he didn't exclusively write horror. In fact, before starting on what was to become his first published novel, *Carrie* (1974), he had already completed four other novels – none of them horror. King described in detail when he wrote those four pre-*Carrie* books, and how he tried, unsuccessfully, to get them published. He does not mention their titles, but refers to them as "Book #1", "#2", and so on. It is understandable that an author would rather remain vague about unpublished material. But in Stephen King's case, he had no choice but to be vague because two of those four books *had* by then been

published under a secret pseudonym: Richard Bachman.

Up to the present-day, King has published no less than seven novels under the pen name Richard Bachman: *Rage* (1977), *The Long Walk* (1979), *Roadwork* (1981), *The Running Man* (1982), *Thinner* (1984), *The Regulators* (1995), and *Blaze* (2007). The uncovering of Bachman's identity sparked shortly after *Thinner* came out. Both readers and reviewers noticed how similar the text was to King novels in style, theme, and narrative drive. In early 1985, King was forced to admit to *The Bangor Daily News* that he was Bachman (Smith 1985). He later acknowledged that it was inevitable that Bachman's true identity would come to light at some point; he had been getting letters from the publication of the very first Bachman book, which he attributed to readers recognizing his "voice" (King 1985, v–vii).

In this article, we will use techniques from computational author identification to determine – albeit, only retrospectively – whether a stylometric analysis of the Bachman books could indeed have discovered Stephen King's distinctive voice in the texts. In stylometry, attribution and verification are commonly distinguished. Authorship attribution is the process of attributing an unknown document to an author within a set of candidate authors (Koppel et al. 2011). Authorship attribution is distinct from authorship verification in that authorship verification involves comparing an unknown text to a corpus of texts by a known author, where the aim is to determine whether the unknown document is also by that author. In authorship verification, no other candidate authors are necessarily involved, although they might serve as a point of comparison (Potha and Stamatatos 2014).

This paper is structured as follows: Below, we will first survey the seminal critical response to the Bachman novels. The next sections present related work and describe the materials used for this study, i.e., the (control) authors and the books selected for analysis. We go on to introduce the experimental setup that we chose for the analysis of the Bachman books and motivate our choice for the Juola protocol (Juola 2015). After discussing the results of our analysis, we move on to the issue of explainability, which is rarely discussed in the context of verification. We describe an experimental method to analyze counts of brand names and pop-culture references in the corpus, one of the "Bachmanesque" features that were explicitly mentioned by early critics.

2. Early Criticism of Bachman Novels

Rage, *The Long Walk*, *Roadwork*, and *The Running Man* appeared as paperback originals with very small print runs. They were hardly reviewed at all. We were only able to find two reviews of *Roadwork* (Slotek 1981; Strachan 1981), and one of *The Running Man* (Frank 1982) – none of them compared Bachman to King. Sam Frank praised Bachman's "vivid gut-level prose," his "straight-ahead storyline", "multi-dimensional characters", and "stark, dank, volcanic descriptions" (1982, 6). Jim Slotek's review of *Roadwork* stated: "Bachman's frenetic, staccato writing style and his habit of throwing in one pop-culture backdrop after another (The Rolling Stones, Merv Griffin, the ever-present television) grab the reader's eye like headlines" (Slotek 1981, 10). That this reviewer noticed enough pop-culture references in the novel to give them such attention in a short review is interesting, since the use of brand names and names of musicians, actors,

songs, movies, and television programs was a “sin” that critics commonly attributed to King. For instance, one negative review of *The Talisman* (1984), which King wrote in collaboration with Peter Straub, posited that the novel “inherited the worst traits of both its parents. From King it has acquired, among other things, his compulsion to list brand-name products, his affinity for pop-cult teenage junk and his penchant for the endless repetition of cryptic italicized phrases” (Lehmann-Haupt 1984, C–15). Reviewer Roger Grooms drew the same conclusion: “Combined novel has King’s flavor. [...] When the frost is on the ‘punkin’, the ubiquitous Stephen King pops up with yet another tale of gruesome goings-on, generally taking place right in the middle of our pop culture” (Grooms 1984, E5).

Around the same time as the publication of *The Talisman* (fall of 1984), Richard Bachman’s *Thinner* arrived in bookstores. The book received more critical attention than the previous four Bachman books because it was published in hardback and review copies were widely distributed. While some reviews made no mention of King (e.g.: O’Neil 1984; Levin 1984; Williams 1984), and one reviewer remarked that the plot of *Thinner* contained “several gaps that a writer like King doesn’t fall through” (Denger 1985, 6D), there were also many reviewers, booksellers, and early readers who heard King’s voice in *Thinner*. One review included the tentative remark that “Bachman’s style is remindful of Stephen King” (Anonymous 1984, 4G), but others were less cautious. *Locus Magazine* wrote: “King does not acknowledge it, but this horror novel about dieting sure sounds like him” (as quoted in Ganley 1985, 5). W. Paul Ganley elaborated on this quote in his own review: “[T]his novel is pure King in style, syntax, character development, dialog, plot structure, humor, gross outs, and even in technical mistakes” (Ganley 1985, 5). Mark Graham announced that “this thriller raises an authorship question”: “If King didn’t write the end of this little narrative, his doppelganger did” (Graham 1984, 26-N).

It is striking that so many readers – independently from one another – seem to have picked up on similarities between Bachman’s and King’s writing style: Out of curiosity alone, this observation sufficiently motivates the question whether computational methods would have been able to pick on these resemblances too. Using an established method from authorship verification, we shall assess below whether that is indeed the case. Perhaps equally striking, however, is that the reviewers’ comments that we have assembled remain rather vague as to why *precisely* they connected Bachman to King; in fact, we could only find a single textual feature mentioned which was concrete enough to be counted using standard digital text analysis techniques: the use of brand names and other references to popular culture. Therefore, we will also explore whether King’s supposed trademark use of this device can be used to single him out as the most-likely author of the Bachman books. As such, this paper contributes to the issue of post hoc explainability in author identification for literary texts.

3. Related Work

Stephen King’s style has received only limited scholarly attention so far. James Arthur Anderson has published two monograph-length studies on King in which he uses linguistic analysis (Anderson 2017, 2020). In *The Linguistics of Stephen King* he examines King’s works “using critical theory developed from linguistic studies” to “provide

a close reading of how King layers language upon language to create both reality and meaning" (Anderson 2017, 7). David L. Hoover dedicated a chapter to King in his monograph *Modes of Composition and the Durability of Style*, examining possible changes in King's style based on chronological periods, changes in mode of composition (handwritten, typewritten, word processed novels), and the difference between writing under the influence and sober. He encountered many categorization problems and could not computationally detect any noticeable stylistic differentiation, concluding that "Stephen's style was not only durable in the face of changes in mode of composition, it was also durable in the face of his abuse of alcohol and drugs and his recovery from that abuse" (Hoover 2021, 165). Van Cranenburgh and Ketzan (2021) also apply computational stylometry to the work of Stephen King to quantify the literariness of 73 King novels and novellas. In their conclusion they suggest that "an exploration of King/Bachman would merit a dedicated, mixed-method study", to determine "whether the Bachman novels (some separated by decades) can convincingly be argued to share distinctive features" and if Bachman is "a signal, or merely noise" (Van Cranenburgh and Ketzan 2021, 196). We propose our paper as a first step in such an exploration.

With the emergence of computing technology in Humanities scholarship, quantitative authorship studies initially focused primarily on two approaches: (1) unsupervised methods (such as exploratory visualization techniques, such as dendrograms and PCA scatterplots), which are still very popular in computational literary studies; or (2) casted the problem of author identification as a standard text classification problem. The latter approach casts attribution as a machine learning task, where exactly one label, from a set of (mutually exclusive) labels, has to be assigned to a previously unseen document. In such a setup, a text classifier can be trained on a set of reference documents for which the authorship can be established beyond reasonable doubt. This particular setup is often referred to as "closed-set attribution", because the set of candidate authors is well delineated and fixed. Excellent performance has been reported in this area, although there still exist important limitations regarding text length, text variety (genre), as well as the number of author classes to be learned.

In many authorship attribution experiments, a text's authorship is determined by calculating the similarity of that text to texts by a set of candidate authors, i.e. a form of "lazy learning". In Burrows's "Delta" method, for example, a text's authorship is predicted based on the frequencies of the frequently-occurring 150 words in the entire set of reference texts (Burrows 2002). "Delta" is a measure that quantifies the similarity between a target text and texts written by potential authors. It represents the standardized difference between the observed frequencies of the 150 most frequently occurring words in the target text and the expected frequencies based on reference texts (Evert et al. 2017). Similarity-based authorship attribution, combined with machine learning methods, have been found to perform well in identifying the true authors of novels written by pseudonymous authors. For example, Jaques Savoy used Burrows's Delta, along with Labbé's distance, nearest shrunken centroids (NSC), naïve Bayes, k-nearest neighbors, and character n-grams to identify Domenico Starnone as the probable author of pseudonymous Italian novelist Elena Ferrante's novels (2018). Eder, Tuzzi, and Cortelazzo corroborate Savoy's prediction that Starnone wrote Ferrante's novels (Eder 2018, Tuzzi and Cortelazzo 2018).

In the real world, however, there are many practical scenarios that do not fit the attribution scenario ideally, mainly because the set of potential candidate authors might be prohibitively large (the “needle in a haystack problem”) and, consequently, it might be difficult (or even impossible) to construct a training data set that can be guaranteed to include the true author of an anonymous document. In such cases, it is problematic that text classifiers will always attribute an anonymous document, no matter what, to one of the available authors, even if none of the available authorial labels in reality applies. Koppel and colleagues have published seminal papers in this area, highlighting that open-set attribution is a much more difficult, but also much more realistic approach to author identification “in the wild” (Koppel and Y. Winter 2014a; Koppel et al. 2007; Koppel et al. 2009). Especially open-set attribution, or authorship verification, has emerged as an established formulation of the problem. Here, algorithms still work with a limited set of candidate authors in the foreground, but specifically take into account the possibility that the correct author might not be included, effectively introducing a “back-off option” where the system returns “None of the above”.

Apart from a series of applied case studies in literary studies, the authorship track in the annual shared task at the PAN workshop has played a major role in benchmarking existing approaches in this domain.¹ Below, we will especially draw inspiration from the “imposter approach” that was seminaly introduced by Koppel and colleagues. Variations of this approach have ranked particularly high in recent editions of the shared task on authorship attribution at PAN. Importantly, the imposter approach is dependent on a pool of “imposter authors” or “distractors”, to which anonymous documents and candidate authors (e.g. in a foreground corpus) can be compared. Researchers have observed that a larger and more diverse imposter pool is invariably beneficial to the performance of the imposter approach. Unsurprisingly, the most successful applications of the method have been applied for text varieties that were abundant and easy to collect online, such as blog posts. For many text varieties, however, it is much more difficult to collect large imposter sets, such as contemporary fiction, because of intellectual rights or digitization backlog.

This is a practical disadvantage of the method that is hard to circumvent in practice. Interestingly, this limitation is bypassed in the so-called Juola protocol that nevertheless still shares characteristics with the imposters approach. Apart from this practical advantage, Juola’s seminal case study (perhaps the most mediatized in the history of the field) bears important resemblances to the Bachman case. After initial speculation in the (social or traditional) media, the high-profile author was twice relatively quick to self-identify as the author behind the pseudonymously published novels. In both cases, moreover, the authors have no reason to consciously alter their writing styles and lacked a clear incentive to publish in an alternative “mode”. This was not the case, for instance, in the well-known French controversy surrounding Romain Gary, where the author actively resisted self-attribution, even after having been called out (Tirvengadam 1996). In the case of Rowling, however, the initial speculation was not based on stylistic similarities, which was the case for Bachman.

Patrick Juola was actively involved in the verification of the authorship of *The Cuckoo’s Calling*, published under the pen name “Robert Galbraith” (Juola 2013a). This successful

1. See: <https://pan.webis.de/> and the overview papers listed there (e.g. Bevendorff et al. 2021).

research initiative later led him to publish a so-called “protocol” (Juola 2015) that included methodological guidelines as to how such cases could be reliably tackled in the future. Juola compared Galbraith’s novel to books by (the small number of) contemporary British female crime novelists (ibid.), that served as distractors. He ultimately operated in an open-set context, because there was no guarantee that the correct author was included among the candidates. As with our Bachman case, sampling a larger pool of imposters was infeasible, because of the intellectual rights that lie on contemporary literature and also often challenging to obtain in a digital format. Thus, the Juola protocol does not implement any sampling of imposters, although it does engage in a stochastic component in the form of feature sampling: A large and diverse feature set is engineered for each of the documents involved and the similarity is measured across these sets to produce a ranking of candidate authors that stable across different feature sets. This approach is reminiscent to the iterative bootstrapping of features in the imposter approach.

Surveying the state of the art in author identification is challenging, because case studies and benchmark task differ enormously, for instance across languages, historical periods, dataset sizes, documents lengths, text varieties involved or the number of candidate authors. Recently, neural models (e.g. Boenninghoff et al. 2019), in particular large foundation models in the form of embedders such as BERT, have yielded promising results. Currently, transformer models appear to be among the best performing author identification method when (1) one is dealing for a language variety for which pretrained language models are available (2) there is a substantial amount of text data available per author. A important survey of modern authorship attribution methods found that performance heavily depends on the number of available words per author in a dataset (Tyo et al. 2022). Tyo et. al found that experiments with datasets containing less than 100,000 words per author, traditional n-gram based models achieved a higher accuracy than BERT-based models (76.50% and 66.71%, respectively). They note that, in general, in experiments with fewer words per author, traditional word- and character-level n-grams outperform more sophisticated deep learning techniques. Their observation is supported across the literature – for example, Alkatori et al. found that using word- and character-level n-grams yield higher accuracy than transformer models in an authorship attribution task using a dataset of *Guardian* articles, with an average of 41 thousand words per author in their corpus. (Altakrori et al. 2021). Thus, while neural methods are promising, they often come with requirements that cannot always be met.

While BERT-based models achieve state-of-the-art accuracy on authorship attribution tasks with extensive word datasets per author, applying such models in the humanities presents challenges due to the technical complexity of deep learning, which may be unfamiliar to researchers in this field. Moreover, model inspection and feature analysis are even more challenging with deep models like BERT compared to traditional n-gram approaches. n-gram models offer simplicity and transparency, making them more suitable for authorship attribution tasks in the humanities where interpretability is crucial. Therefore, in this paper, we opt for a simple similarity-based authorship attribution method that relies on transparent word- and char-level n-gram document representations.

4. Materials

Our corpus of distractor novels is made up of texts by three horror-thriller writers: Dean Koontz, Peter Straub, and Thomas Harris. These writers were chosen because, like King, all three are American, male authors that published popular novels in the 1970s, 80s, 90s, and 2000s in the same genre. The corpus includes 20 novels by King, 5 by Harris, 12 by Straub, and 17 by Koontz. It consists of all novels by Harris and Straub and a selection of books by Koontz and King up until 2007, which is when *Blaze* was published. The books were obtained in EPUB format and converted to UTF-8-encoded plain text files. [Table 1](#) includes a selection of lexical statistics for each of the texts. The number of unique tokens, i.e., types, may be affected by text length: Type-token ratios tend to be lower in longer texts because words are more likely to reoccur in longer texts (Richards 1987). Therefore, type-token ratios (TTR) were extracted from the first 10,000 tokens of each book ([Table 1](#)).

5. Task Operationalization

In this section, we investigate whether it is possible to identify Stephen King, only post hoc of course, as the author of the Bachman books. Our approach reproduces some of the key characteristics of Juola's authorship verification protocol, who recommends building various feature sets (word lengths, most frequent words, character 4-grams, and word bigrams) to calculate similarities between the target and distractor texts (Juola 2013b). The author who wrote the text most similar to a target text is predicted to be the author of the target text.

As in Juola's protocol, we convert calculated similarities to ranks. Juola takes similarities between target texts and known-author texts and ranks each author by their comparative similarity to the target text. Likewise, we take the cosine distances between book segments, i.e., sequences of consecutive tokens drawn from a book. We calculate the cosine distance between a Bachman segment and a segment by one of the candidate authors in our corpus, and rank the candidate authors by cosine distance. Consider a Bachman segment that has the cosine distances 0.43 for a Harris segment, 0.30 for a King segment, 0.70 for a Koontz segment, and 0.67 for a Straub segment. For this Bachman segment, King occupies rank 1 because the King segment had the smallest cosine distance to the Bachman segment, Harris has rank 2 because his segment was second closest, and so on.

Our prediction algorithm is based on Koppel and Winter's many-candidates method of authorship attribution (Koppel and Y. Winter 2014b). It relies on several varied ("bootstrapped") feature sets used to make predictions. Using diverse feature sets to represent texts reduces false similarities between target texts and known-author texts that cannot be reproduced with different feature sets. Half of the features from this varied feature set are randomly subsampled to calculate the similarity between a target text and a known-author text, repeated in k iterations. Finally, candidate authors are each assigned a score representing the proportion of iterations in which the candidate's segment was most similar to the target segment.

Author	Title	Date of Publication	Word Count	Number of Word Types	TTR (First 10,000 Tokens)
Bachman	1966	<i>The Long Walk</i>	87,333	7,928	0.197
	1968	<i>Roadwork</i>	93,047	8,786	0.209
	1970	<i>Rage</i>	55,909	6,203	0.209
	1973	<i>Blaze</i>	82,444	7,720	0.187
	1981	<i>The Running Man</i>	67,769	8,579	0.252
	1984	<i>Thinner</i>	99,272	8,582	0.221
	1995	<i>The Regulators</i>	120,909	9,456	0.217
Harris	1975	<i>Black Sunday</i>	96,485	9,194	0.247
	1981	<i>Red Dragon</i>	105,648	9,136	0.214
	1988	<i>The Silence of the Lambs</i>	99,299	8,878	0.222
	1999	<i>Hannibal</i>	126,831	11,588	0.240
	2006	<i>Hannibal Rising</i>	67,575	7,671	0.231
King	1974	<i>Carrie</i>	62,275	7,509	0.247
	1975	<i>'Salem's Lot</i>	156,566	12,117	0.251
	1977	<i>The Shining</i>	165,734	11,574	0.218
	1978	<i>The Stand</i>	479,256	20,363	0.217
	1979	<i>The Dead Zone</i>	156,648	11,689	0.246
	1994	<i>Insomnia</i>	251,490	13,842	0.215
	1995	<i>Rose Madder</i>	180,040	11,171	0.205
	1996	<i>The Green Mile</i>	135,954	8,754	0.212
	1996	<i>Desperation</i>	199,619	11,581	0.202
	1997	<i>Wizard and Glass</i>	265,321	13,470	0.209
	1998	<i>Bag of Bones</i>	215,488	13,045	0.214
	1999	<i>The Girl Who Loved Tom Gordon</i>	63,368	6,102	0.201
	2001	<i>Dreamcatcher</i>	214,223	13,051	0.195
	2002	<i>From a Buick 8</i>	128,836	9,159	0.195
	2003	<i>Wolves of the Calla</i>	251,658	13,232	0.199
	2004	<i>Song of Susannah</i>	132,659	10,018	0.198
	2004	<i>The Dark Tower</i>	283,647	14,861	0.214
	2005	<i>Colorado Kid</i>	35,265	4,260	0.201
	2006	<i>Cell</i>	126,858	9,235	0.208
		2006	<i>Lisey's Story</i>	192,276	11,719
Koontz	1968	<i>Star Quest</i>	35,233	5,057	0.256
	1970	<i>Beastchild</i>	48,544	5,900	0.224
	1972	<i>Warlock</i>	57,259	6,825	0.240
	1974	<i>After the Last Race</i>	84,411	7,895	0.221
	1976	<i>Night Chills</i>	94,934	8,477	0.241
	1977	<i>The Vision</i>	66,100	6,508	0.215
	1980	<i>Whispers</i>	177,987	11,509	0.246
	1981	<i>The Eyes of Darkness</i>	89,325	8,299	0.245
	1983	<i>Phantoms</i>	138,378	10,878	0.225
	1986	<i>Strangers</i>	264,107	15,553	0.270
	1988	<i>Lightning</i>	140,257	10,676	0.234
	1990	<i>The Bad Place</i>	147,579	11,311	0.248
	1992	<i>Hideaway</i>	130,796	11,235	0.253
	1994	<i>Winter Moon</i>	118,019	10,376	0.250
	1998	<i>Fear Nothing</i>	130,790	11,553	0.249
	2000	<i>From the Corner of His Eye</i>	217,821	15,649	0.267
	2002	<i>By the Light of the Moon</i>	127,314	11,920	0.302
	2004	<i>The Taking</i>	86,104	9,857	0.280
	2006	<i>The Husband</i>	86,316	8,797	0.231
		2007	<i>The Good Guy</i>	84,424	8,305
Straub	1975	<i>Julia</i>	85,342	7,853	0.211
	1977	<i>If You Could See Me Now</i>	110,624	8,394	0.236
	1979	<i>Ghost Story</i>	187,951	11,283	0.217
	1980	<i>Shadowland</i>	159,291	10,621	0.219
	1982	<i>Floating Dragon</i>	225,917	13,022	0.252
	1988	<i>Koko</i>	210,205	12,503	0.238
	1990	<i>Mystery</i>	181,264	10,243	0.221
	1993	<i>The Throat</i>	254,781	12,635	0.218
	1995	<i>Hellfire Club</i>	191,789	11,760	0.232
	1999	<i>Mr. X</i>	186,987	13,010	0.244
	2003	<i>Lost Boy, Lost Girl</i>	89,152	8,044	0.244
	2004	<i>In The Night Room</i>	103,235	9,070	0.232

Table 1: Basic statistics about the books used in this corpus. Included are the total number of tokens (word count), unique tokens (word types), and the type-token ratio (TTR) in the first 10,000 tokens.

First, we produce a corpus of segments. To create this corpus, we tokenize and remove punctuation from all books in our corpus. Each book is then split into segments of 1,000, 5,000, and 10,000 consecutive tokens. Trailing segments of less than each of the aforementioned lengths are not included in the dataset.

Second, we produce a large feature set from the segments. We generate the feature set by vectorizing segments using combinations of different vectorizer settings. Half of the vectorizers use tf-idf weighting, and half of the vectorizers do not use tf-idf weighting. We use char and word analyzers with ranges of 2 to 4, and 1 to 3, respectively; i.e., word vectorizers create features of unigrams, bigrams, and trigrams, and char vectorizers create features of bigrams, trigrams, and 4-grams. Each vectorizer caps the number of features/columns it extracts at 10,000 to limit the number of n-grams that only appear in one book.

In total, the combination of 2 tf-idf settings (true and false), 2 analyzers (word and char), and n-gram ranges of 3 (for both word and char n-grams) creates 12 distinct vectorizer settings to generate 12 feature spaces. The 12 feature spaces are concatenated to create one combined feature space with a maximum of 120,000 columns. Each new feature space was scaled with min-max scaling before being appended to the final feature space.

Third, we calculate the cosine distance between each Bachman segment vector and a random segment vector by each distractor author.

Below is our algorithm in pseudo-code:

For each segment length (1,000, 5,000, and 10,000 tokens) s :

1. Split all books in the corpus into segments of s tokens;
2. Initialize an empty list of feature spaces l ;
3. For each collection of vectorizer settings v :
 - (a) Instantiate a vectorizer i with collection of settings v ;
 - (b) Create a feature space f by vectorizing the corpus of segments using vectorizer i ;
 - (c) Append the feature space f in the list of feature spaces l ;
4. Horizontally concatenate the list of feature spaces l into a 2-dimensional feature space array a ;
5. For each row r_B representing a Bachman segment in the feature space array a , repeat 1,000 times:
 - (a) For each candidate author in the corpus (King, Koontz, Straub, and Harris):
 - i. Randomly select a row r_C representing a segment by this candidate author in the feature space a ;
 - ii. Randomly sample 10,000 distinct features from the feature space a ;
 - iii. Calculate the cosine distance between r_B and r_C using these 10,000 randomly-chosen features;

Title	Rank	Harris	King	Koontz	Straub
<i>The Long Walk</i>	1	0.018	0.661	0.077	0.244
	2	0.101	0.238	0.214	0.446
	3	0.311	0.080	0.386	0.223
	4	0.569	0.021	0.323	0.087
<i>Roadwork</i>	1	0.079	0.491	0.108	0.322
	2	0.181	0.290	0.183	0.346
	3	0.322	0.156	0.303	0.219
	4	0.418	0.063	0.406	0.114
<i>Rage</i>	1	0.043	0.526	0.062	0.370
	2	0.141	0.324	0.158	0.377
	3	0.342	0.117	0.362	0.178
	4	0.474	0.033	0.418	0.075
<i>Blaze</i>	1	0.055	0.616	0.085	0.244
	2	0.184	0.244	0.196	0.376
	3	0.354	0.102	0.305	0.239
	4	0.407	0.038	0.414	0.140
<i>The Running Man</i>	1	0.086	0.456	0.165	0.293
	2	0.193	0.267	0.220	0.321
	3	0.309	0.172	0.281	0.238
	4	0.413	0.105	0.334	0.148
<i>Thinner</i>	1	0.032	0.604	0.104	0.260
	2	0.117	0.262	0.229	0.392
	3	0.296	0.103	0.363	0.239
	4	0.555	0.032	0.304	0.109
<i>The Regulators</i>	1	0.022	0.706	0.085	0.187
	2	0.115	0.203	0.250	0.432
	3	0.296	0.069	0.373	0.262
	4	0.567	0.021	0.293	0.118

Table 2: Bachman book titles, ranks, and the proportions of 5,000-token segments from each book that were predicted to be written by each author in the corpus. Tables for data collected from 1,000- and 10,000-token segments can be viewed in the Appendix (Table 4 and Table 5).

- Convert cosine distances into ranks (1 = segment has lowest cosine distance to Bachman segment, 4 = segment has highest cosine distance to Bachman segment).

This algorithm created 1,000 cosine distances between each Bachman segment and a segment by a candidate author – 4,000 cosine distances in total per Bachman segment. We describe the results of this analysis below.

6. Authorship Attribution Results

In order to test whether authors received certain rankings significantly more or less often than if rankings were sampled from a random distribution, we performed a chi-squared test and visualized Pearson residuals using an association plot (Figure 1). The association plots show how ranking counts vary in a meaningful way across authors. They indicate that King was considerably more likely to be predicted as the author of a Bachman segment (ranking = 1). The other candidate authors were significantly less likely to be predicted as the author of a Bachman segment. In plots created using 5,000 and 10,000-token segments, segments written by King were significantly less likely to be 2nd, 3rd, or 4th most similar to Bachman segments (Figure 2). However, in the plot created using 1,000-token segments, segments written by King were significantly more

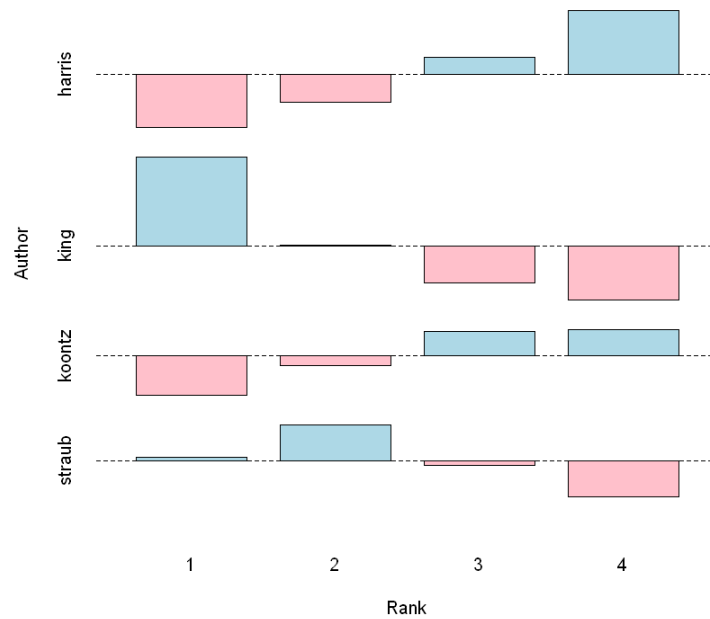


Figure 1: Direction and significance of correlation between the number of times a candidate author received a similarity ranking (1, 2, 3, or 4) to Bachman segments and candidate author in 5,000-token segments. See Figure 2 for plots of rankings in 1,000- and 10,000- token segments.

likely to be both the 1st and 2nd most similar to Bachman segments.

For every Bachman book, Stephen King was predicted as the author of a segment in a much greater proportion of iterations than all other candidate authors (Table 2). *The Regulators* had the highest proportion of iterations in which King was the predicted author, at 50.8%, 70.6%, and 81.8% of iterations with 1,000, 5,000, and 10,000-token segments. By contrast, *The Running Man* had the lowest proportion of iterations in which King was the predicted author for 1,000, 5,000 and 10,000-token segments, with 36.8%, 45.6%, and 47.5%, respectively. King nevertheless still received the highest proportion of iterations compared to other authors.

7. Brand Name and Pop-Culture References Analysis

King’s work has been described as combining the tradition of American naturalism with the classic supernatural horror genre (Bradley 1998, 96). As he himself has vehemently stated, King was in no way the first writer to take horror out of its classic gothic settings and transport it into small-town America. He has claimed that Richard Matheson, Robert Bloch, Jack Finney, and the TV show *The Twilight Zone* created the genre of modern American horror, which lies at the root of his poetics: “Those things formed my idea of what a horror story should do: The monster shouldn’t be in a graveyard in decadent old Europe, but in the house down the street” (Underwood and Miller 1989, 93). The craft, in King’s opinion, is to “create any kind of environment that the reader can identify

with totally" (Thomases and Tebbel 1981, 95) and then to "inject [it] with the fantasy element" (Underwood and Miller 1989, 113).

Inserting brand names and references to pop-culture is a tool in creating the familiarity necessary for the optimal effect of the horrific. In the essay "Dean Koontz and Stephen King: Style, *Invasion*, and an Aesthetics of Horror", Michael R. Collings posits that "brand-name descriptions, carefully established realism of setting and character, common images, and themes may themselves become, not trademarks of a single author (as we frequently assume when we talk of King's brand names), but characteristics of dark fantasy itself, part of the realism of presentation that C. S. Lewis argued was essential to fantasy at any level" (Collings 1998, 76). Including brand names in their work is a practice King and Koontz share, Collings believes, because it is inherent in the genre, "although far more so in King than in Koontz" (Collings 1998, 76). Thus, in our opinion, it would be more interesting to approach the use of brand names and pop-culture references as a genre convention than as an idiosyncrasy of King's style, and to test whether King does indeed use brand names and references to pop-culture significantly more often than other authors in the genre – so much so that attentive readers could have deduced that the brand names used by Bachman "sounded" like King.

As a first step, we aimed to quantify the references to brand names and pop-culture of all five authors. We manually compiled lists of such references in books from the authors in our corpus: Four of the five Bachman books published before King was uncovered as the author (*Rage*, *The Long Walk*, *Roadwork*, and *Thinner*),² and three novels by Koontz, Straub, King, and Harris. We extracted references from three texts by each to avoid the risk of choosing one novel that may not accurately represent the author's overall writing style. This allowed us to establish an average use of references for all authors. Where possible, we selected novels that were published during the same period as the early Bachman books: between 1977 and 1984.³ These were the novels that fans of the genre of modern horror would have read not long before *Thinner* came out.

The concepts "popular culture" and "brand name" are difficult to define and delineate, so we have opted to cast a wide net: We include not only names of contemporary celebrities (e.g., "Kitty Carlisle", "Sting"), brand names of commercial products (e.g., "Ford", "Shell"), TV shows ("I Love Lucy", "Star Trek"), and movies ("Wizard of Oz", "Psycho"), but also the names of newspapers, hotels, airlines, sports teams, banks, musicians, writers, painters, literary characters, and book titles. Each reference was counted only once, no matter how many times it occurs in the text. In King's *Cujo*, for instance, a Ford Pinto plays an important role, but it only counts as one of the 235 references found in the novel. The result is presented in Table 3, and the complete lists of extracted references for each author can be viewed in subsection A.2.

To enable comparison, the number of references in each text was normalized to a standard rate per 100,000 words. We then calculated the average use of references for the total of all novels per author. As indicated in the table, Bachman included an average

2. Since *The Running Man* is set in a then-distant future, it hardly contains such references, so we disregarded it from this experiment.

3. In the case of Thomas Harris, who only published one novel in our chosen time period, we used *Black Sunday* (1975), *Red Dragon* (1981) and *The Silence of the Lambs* (1988).

Author	Title	References	Word Count	Refs per 100,000 words
King	<i>Firestarter</i>	202	153,219	132
	<i>Cujo</i>	235	119,497	196
	<i>Pet Sematary</i>	215	144,961	148
	average:			168
Koontz	<i>The Eyes of Darkness</i>	60	89,325	68
	<i>Phantoms</i>	93	135,058	68
	<i>Darkfall</i>	63	102,550	62
	average:			66
Straub	<i>Ghost Story</i>	212	182,732	116
	<i>Shadowland</i>	115	159,291	72
	<i>Floating Dragon</i>	137	225,917	60
	average:			82
Harris	<i>Black Sunday</i>	114	96,485	118
	<i>Red Dragon</i>	87	105,648	82
	<i>The Silence of the Lambs</i>	142	99,299	144
	average:			114
Bachman	<i>Rage</i>	141	55,909	252
	<i>The Long Walk</i>	55	87,561	62
	<i>Roadwork</i>	225	93,272	242
	<i>Thinner</i>	218	99,272	220
	average:			190

Table 3: Unique references to brand names and pop-culture in 3 books by King, Koontz, Straub and Harris, and in 4 books by Bachman.

of **190** unique references per 100,000 words, followed by King with **168**, Harris with **114**, Straub with **82**, and Koontz with **66**.

In the second phase of our analysis, we sought to determine whether there was a significant overlap between the brand names and pop-culture references used by King in his Bachman books and those used in the novels published under his own name. This overlap could potentially reveal King's "voice" in the Bachman books through his selection of references.⁴ To test this, we examined how many of the 517 references found in Bachman's works also appeared in the texts of the other authors.⁵ All Bachman, King, Koontz, Straub, and Harris books were analyzed using a software library (*SpaCy*) capable of automatically tagging named entities, including those consisting of multiple tokens (e.g., "I Love Lucy"). Our algorithm is as follows:

1. For all books (Bachman, King, Koontz, and Harris books), repeat 100 times:
 - (a) Initialize a total cultural reference count of 0;
 - (b) Randomly select a 10,000-token segment in the book;
 - i. For each manually-collected pop-culture reference;
 - A. Count the number of SpaCy-extracted named entities whose text matches the manually collected pop-culture reference and add it to the total cultural reference count;

4. Having such a distinctive stylistic trait as an author also creates the risk that an imposter could insert similar references into a story to falsely present it as a King novel. Nonetheless, within the scope of this test case, the ease with which this trait can be imitated does not pose a problem.

5. We removed 8 items from the list of pop-culture references found in Bachman's works that are names of characters in other books in the corpus and would therefore inflate counts of pop-culture references in these books. The list of 8 excluded pop-culture references and the reason why each was excluded can be found in the Appendix (Table 6).

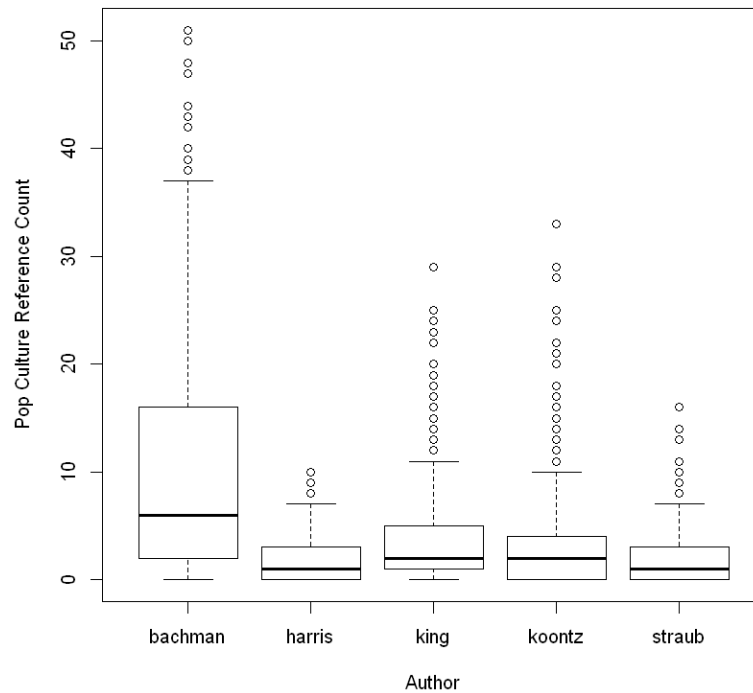


Figure 2: Boxplot showing the absolute frequencies of pop-culture references in 100 randomly-selected 10,000-token segments from each Bachman, Harris, King, Koontz, and Straub book.

- (c) Store the total cultural reference count in a list of cultural reference counts for each book.

Our algorithm produces 100 pop-culture reference counts per book in the corpus – 64,000 counts in total. Differences in the central tendencies of these pop-culture reference counts by author are compared in a pairwise fashion with a Wilcoxon rank-sums test. Wilcoxon rank-sums tests compare pop-culture reference counts in Bachman versus King books, Bachman versus Koontz books, Bachman versus Harris books, and Bachman versus Straub books. In addition, Wilcoxon rank-sums tests compare pop-culture reference counts in King versus Koontz books, King versus Harris books, and King versus Straub books.

The cultural references extracted from Bachman books were, as could be expected, significantly more common in Bachman segments than in segments by any other author in the corpus. A one-tailed two-sample Wilcoxon rank sums test indicated that the median pop-culture reference count was significantly higher in Bachman segments than in King segments ($W = 963,633, p < 0.001$) (Table 7). Median pop-culture reference counts in Bachman segments were also found to be significantly higher than in segments by Harris, ($W = 274,381, p < 0.001$), Koontz ($W = 1,011,395, p < 0.001$), and Straub ($W = 662,439, p < 0.001$). That is to be expected, however, since the reference list was compiled on the basis of these very texts.

Of the distractor authors, King segments contained the most references extracted from Bachman books. One-tailed two-sample Wilcoxon rank-sums tests were applied to each pair of distractor authors – King and Straub, King and Harris, King and Koontz

(Table 8). Median pop-culture reference counts in King segments were found to be significantly higher than in segments by Straub ($W = 1,560,287, p < 0.001$), Koontz ($W = 2,222,412, p < 0.001$), and Harris ($W = 639,999, p < 0.001$). Table 9 contains a breakdown of pop-culture reference counts by book title.

8. Discussion

The results of our analysis suggest that computational methods can correctly identify King as the real author of the Bachman books. However, the chosen segment length matters – larger segments lengths seem to produce more extreme, and probably less trustworthy proportions for any given rank and author. The proportions of iterations in which a Bachman segment was predicted to be written by King increased with segment length. Likewise, authors that were more consistently ranked third or fourth, like Dean Koontz and Thomas Harris, had higher proportions of iterations with rank 3 or 4 in longer segment lengths. This trend is consistent with the observation that larger text sizes (5,000 tokens and over) tend to increase the probability that a text's authorship will be correctly attributed (Eder 2015). The extreme proportions in 10,000-token segments are likely the result of skew by a smaller sample size.

While *Thinner* led to readers outing King as the true author of the Bachman books, *Thinner* did not have the highest proportion of iterations predicting King as the author of its segments, *The Long Walk* did, a text written so early in his career that it could be classified under juvenilia. It was *Thinner* that led to King's unmasking, not because it was more "King-like" than the previous Bachman books, but because it had a larger readership and was much more widely reviewed. *The Regulators*, the only novel conceived specifically for the alter ego "Richard Bachman" – an alter ego that had surely taken form in King's mind by the mid-nineties – scored second-lowest (scoring only fractionally better than *The Running Man*), which in our opinion is mainly due to the different text-types which take up a substantial portion of the novel: long letters and sections from a screenplay; but it could also be because of King's conscious effort, as he stated afterwards, to find "a good voice and a valid point of view that were a little different from my own" (King 1996, ix), something he did not do with the previous novels that were only published as Bachman but not written as him. While it has been argued that all Bachman books share certain features (Collings 2011, Strengell 2005), our experiment suggests that stylistically these texts are not noticeably different to his other work, which confirms David L. Hoover's findings regarding the durability of King's style: His literary voice was present (and detectable) from very early on and has not been subject to much change throughout his career. We support Van Cranenburgh and Ketzan's suggestion for further exploration of King/Bachman, to perhaps identify, by computational means, the small difference in voice that King himself hears in his Bachman novels.

Interestingly, Peter Straub consistently has the highest value for rank 2. King and Straub, who were friends and collaborators, have commented on each other's styles over the years. Straub praised the style of *The Shining* as the "reverse" of a literary style which "made a virtue of colloquialism and transparency", an "unprecedentedly direct style" (Straub 1984, 10). Because of this directness and transparency, it has been claimed that

King in fact has “no style” (Bradley 1998, 116). The language, King believes, should not pose an interference between the reader and the story, it should be accessible to all, allowing the reader to “get through the barrier of print and into the story without too much effort”, and for that to happen, the “writer’s voice” should be “low enough” (Dewes 1981, 63). Straub’s style is more classic; referred to as “the good prose” by King, “almost always structurally correct”, “not flashy”, and “unobtrusively strong” (King 1982b, 30). King has noted stylistic differences between Straub’s early novels: He called *Julia* (1975) an “English ghost story”, its diction also being “English – cool, rational, almost disconnected from any kind of emotional base” (King 1982a, 285); *If You Could See Me Now* (1977) has a “Chandleresque first-person narrative” (King 1982b, 31); and *Ghost Story* (1979) – Straub’s third novel in a row to feature a ghost – is written in a “Jamesian diction” (King 1982b, 31). The slightly meandering style in Straub’s early works might in part account for Straub emerging from the experiment as the second-most-likely author of the Bachman books. Of relevance here also is that Straub has stated that King’s early novels *Salem’s Lot* and *The Shining* heavily influenced him as a writer: “[King’s] aims and ambitions were very close to my own [...] [*The Shining*] was like a roadmap of where to go: [King] armored my ambition” (Straub 1984, 9–10). The differences between their styles did not propose a challenge when they collaborated on *The Talisman*. As Straub told an interviewer in 1985: “Our styles seemed to melt together. The book has its own sound; it doesn’t sound like me and it doesn’t sound like Steve. [...] There were times when I deliberately imitated Steve’s style and there were times when he deliberately, playfully, imitated mine” (D. E. Winter 1985, 64). Straub’s rank 2 in our experiment is an indication that a comparative stylistic analysis of the works of King and Straub, including their two collaborations, would be a fruitful avenue of further research.

The results of the pop-culture experiments reveal that King’s use of brand names and pop-culture references was part of his style from the first stages of his career, when he wrote the early Bachman books. *Rage*, in particular, which he began at the age of nineteen, contains a high amount of references: 141 in a 55,909-word text. *Rage* is not a horror novel, but a suspense novel set in contemporary USA, as is *Roadwork*, which has an equally high amount of references: 225 in a 93,272-word text. Converted to averages per 100,000 words, these two books each contain over three times as many references as any of the novels by Koontz and Straub. *Rage* and *Roadwork*, both non-horror books, show that King’s use of brand names and popular culture as a technique to create a world that is familiar to the reader transcended (and in a sense predated) his views on the poetics of the modern horror genre; the real world is there even when the fantasy element is not.

The manual counts of references indicate that all four modern horror practitioners used the technique in their novels. This tentatively confirms the claim that it is inherent to the genre. However, King employed it to a much greater extent than Straub and Koontz, which irked some literary critics in the nineteen eighties. A larger-scale study of this kind, with a corpus consisting of more works by more authors, would provide a more accurate picture of the genre’s dependence on realism and familiarity in setting, for instance in comparison with other genres in popular fiction, such as romances, mystery novels, and thrillers. There is a challenge, however, in the degree to which the extraction of the references can be automated. Reliable named entity recognition is a vital first step

in automating this kind of work. However, a filtering process is necessary to remove or flag irrelevant entries such as character and place names for it to be feasible for all remaining entries to be manually vetted, since the named entities in a novel easily run into the thousands.

9. Conclusion

Style is elusive and idiosyncratic. While some reviewers were confident King was the real author of the Bachman books, they remained vague about the similar stylistic features they had discovered in the texts of both authors. In our paper, we showed that an authorship attribution algorithm is able to predict King to be the author of Bachman texts significantly more often than the distractor authors. But because it predicts authorship by randomly sampling large feature sets in thousands of iterations, the algorithm is a black box, which, comparable to the reviewers, only outputs scores and does not supply insights into the similarity in voice.

Our research might be extended by applying machine learning and deep learning authorship attribution techniques. In this paper, we use a similarity-based authorship attribution method instead of machine learning or deep learning methods. Because pre-trained BERT models have demonstrated state-of-the-art accuracy in authorship attribution experiments where each candidate author in the corpus has a large number of words, we expect BERT-based models to achieve better results with our corpus. It remains very much the question, however, whether such gains in performance would eventually scale to lesser-resourced (e.g. historic) literary cultures.

This paper also explored how pop-culture references may be important for identifying King as the real author of the Bachman books. Brand names and pop-culture references extracted from Bachman books were found to be significantly more common in King books than in books by Koontz, Straub, or Harris. Moreover, *Rage*, *Roadwork* and *Thinner* contained much more references than any of the contemporary novels by Koontz, Straub, and Harris, which again pointed towards King being the author (*Cujo* being similarly packed with pop-culture, for instance). These findings indicate a promising direction for further exploration of this technique as an inherent feature in modern horror fiction, but also shed some light on the intuition of many readers at the time that the Bachman books had actually been written by the author who, as he suggested in his essay in *Adelina*, had become a brand name himself.

10. Data Availability

Due to copyright restrictions, the full texts and segments of the King, Straub, Harris, and Koontz books used in our experiments cannot openly be shared. Data extracted from full texts can be found here: <https://zenodo.org/doi/10.5281/zenodo.7956048>.

11. Software Availability

Software can be found here: https://github.com/dorothyh-ms/king_bachman_authorship_verification.

12. Acknowledgements

Funded in part by the FWO research project "Creating Suspense Across Versions: Genetic Narratology and Stephen King's IT" at the University of Antwerp.

13. Author Contributions

Dorothy Modrall Sperling: Conceptualization, Writing – original draft

Vincent Neyt: Conceptualization, Writing – original draft, Writing – review & editing

Mike Kestemont: Conceptualization, Methodology, Writing – review & editing

References

- Altakrori, Malik, Jackie Chi Kit Cheung, and Benjamin C. M. Fung (2021). "The Topic Confusion Task: A Novel Evaluation Scenario for Authorship Attribution". In: *Findings of the Association for Computational Linguistics: EMNLP 2021*, 4242–4256. [10.18653/v1/2021.findings-emnlp.359](#).
- Anderson, James Arthur (2017). *The Linguistics of Stephen King: Layered Language and Meaning in the Fiction*. McFarland.
- (2020). *Excavating Stephen King: A Darwinist Hermeneutic Study of the Fiction*. Lexington Books.
- Anonymous (1984). "No Dieting after 'Thinner'". In: *The Kingsport Times-News*, 4G.
- Bevendorff, Janek, BERTa Chulvi, Gretel Liz De La Peña Sarracén, Mike Kestemont, Enrique Manjavacas, Ilia Markov, Maximilian Mayerl, Martin Potthast, Francisco Rangel, Paolo Rosso, et al. (2021). "Overview of PAN 2021: Authorship Verification, Profiling Hate Speech Spreaders on Twitter, and Style Change Detection". In: *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 12th International Conference of the CLEF Association, CLEF 2021, Virtual Event, September 21–24, 2021, Proceedings 12*. Springer, 419–431. [10.1007/978-3-030-72240-1_66](#).
- Boenninghoff, Benedikt, Steffen Hessler, Dorothea Kolossa, and Robert M. Nickel (2019). "Explainable Authorship Verification in Social Media via Attention-based Similarity Learning". In: *IEEE International Conference on Big Data (IEEE Big Data 2019)*, Los Angeles, CA, USA, December 9–12, 2019. [10.1109/BigData47090.2019.9005650](#).
- Bradley, Linda (1998). "The Sin Eater: Orality, Postliteracy, and the Early Stephen King". In: *Stephen King*. Ed. by Harold Bloom. Bloom's Modern Critical Views. Chelsea House, 95–124.
- Burrows, John (2002). "'Delta': a Measure of Stylistic Difference and a Guide to Likely Authorship". In: *Literary and Linguistic Computing* 17 (3), 267–287. [10.1093/l1c/17.3.267](#).
- Collings, Michael R. (1998). "Dean Koontz and Stephen King: Style, *Invasion*, and an Aesthetics of Horror". In: *Discovering Dean Koontz: Essays on America's Bestselling Writer of Suspense*. Ed. by Bill Munster. Wildside Press, 64–79.
- (2011). *Stephen King is Richard Bachman*. Overlook Connection Press.
- Denger, Laurie (1985). "Bachman Novel has Thrills, Chills, Gaps". In: *The Dayton Daily News*, 6D.

- Dewes, Joyce Lynch (1981). "Interview: Stephen King". In: *Mystery Magazine*.
- Eder, Maciej (2015). "Does Size Matter? Authorship Attribution, Small Samples, Big Problem". In: *Digital Scholarship in the Humanities* 30 (2), 167–182. [10.1093/llc/fqt066](https://doi.org/10.1093/llc/fqt066).
- (2018). "Elena Ferrante: a Virtual Author". In: *Drawing Elena Ferrante's Profile*. Ed. by Arjuna Tuzzi and Michele Alberto Cortelazzo, 31–46.
- Evert, Stefan, Thomas Proisl, Fotis Jannidis, Isabella Reger, Steffen Pielström, Christof Schöch, and Thorsten Vitt (2017). "Understanding and Explaining Delta Measures for Authorship Attribution". In: *Digital Scholarship in the Humanities* 32 (suppl₂), ii4–ii16. [10.1093/llc/fqx023](https://doi.org/10.1093/llc/fqx023).
- Frank, Sam (1982). "Running Man Beats the Odds". In: *The San Francisco Examiner*, 6.
- Ganley, W. Paul (1985). "Thinner, by Richard Bachman". In: *Fantasy Mongers* 13.
- Graham, Mark (Dec. 1984). "Fit for a King: This Thriller Raises an Authorship Question". In: *The Rocky Mountain News*, 26–N.
- Grooms, Roger (Nov. 1984). "Combined Novel has King's Flavor". In: *Palladium-Item*, E5.
- Hoover, David L. (2021). *Modes of Composition and the Durability of Style in Literature*. Routledge.
- Juola, Patrick (2013a). "How a Computer Program Helped Show J.K. Rowling Write A Cuckoo's Calling". In: *Scientific American*. <https://www.scientificamerican.com/article/how-a-computer-program-helped-show-jk-rowling-write-a-cuckoos-calling/> (visited on 11/29/2023).
- (July 2013b). *Rowling and 'Galbraith': an Authorial Analysis*. Language Log. UPenn. <https://language-log.ldc.upenn.edu/nll/?p=5315> (visited on 11/29/2023).
- (2015). "The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions". In: *Digital Scholarship in the Humanities* 30 (suppl₁), i100–i113. [10.1093/llc/fqv040](https://doi.org/10.1093/llc/fqv040).
- King, Stephen (1982a). *Danse Macabre*. Futura.
- (1982b). "Peter Straub: An Informal Appreciation". In: *Program Book, World Fantasy Convention '82*. Ed. by Kennedy Poyser. World Fantasy Convention.
- (1985). *The Bachman Books*. New American Library.
- (1996). "The Importance of Being Bachman". In: *The Bachman Books*. Plume.
- Koppel, Moshe, Jonathan Schler, and Shlomo Argamon (2009). "Computational Methods in Authorship Attribution". In: *Journal of the American Society for Information Science and Technology* 60 (1), 9–26. [10.1002/asi.20961](https://doi.org/10.1002/asi.20961).
- (2011). "Authorship Attribution in the Wild". In: *Language Resources and Evaluation* 45 (1), 83–94. [10.1007/s10579-009-9111-2](https://doi.org/10.1007/s10579-009-9111-2).
- Koppel, Moshe, Jonathan Schler, and Elisheva Bonchek-Dokow (2007). "Measuring Differentiability: Unmasking Pseudonymous Authors". In: *Journal of Machine Learning Research* 8 (45), 1261–1276. <http://jmlr.org/papers/v8/koppel07a.html> (visited on 11/29/2023).
- Koppel, Moshe and Yaron Winter (2014a). "Determining if Two Documents Are Written by the Same Author". In: *Journal of the Association for Information Science and Technology* 65 (1), 178–187. [10.1002/asi.22954](https://doi.org/10.1002/asi.22954).
- (2014b). "Determining if Two Documents Are Written by the Same Author". In: *Journal of the Association for Information Science and Technology* 65 (1), 178–187. [10.1002/asi.22954](https://doi.org/10.1002/asi.22954).

- Lehmann-Haupt (Nov. 1984). "An Ungainly Offspring". In: *The Sunday Herald-Times*, C-15.
- Levin, Bob (Dec. 1984). "Novel Cursed by Cliches, Thin Characterizations". In: *The Atlanta Constitution*, 9-J.
- O'Neil, Ann W. (Dec. 1984). "A Horrifying Weight-Loss Plan". In: *The Philadelphia Daily News*, 51.
- Potha, Nektaria and Efstathios Stamatatos (2014). "A Profile-Based Method for Authorship Verification". In: *Hellenic Conference on Artificial Intelligence*. Springer, Cham, 313-326. [10.1007/978-3-319-07064-3_25](https://doi.org/10.1007/978-3-319-07064-3_25).
- Richards, Brian (1987). "Type/Token Ratios: What Do They Really Tell Us?" In: *Journal of Child Language* 14 (2), 201-209. [10.1017/S0305000900012885](https://doi.org/10.1017/S0305000900012885).
- Slotek, Jim (June 1981). "Roadwork, by Richard Bachman". In: *The Ottawa Citizen*, 10.
- Smith, Joah H (Feb. 1985). "Pseudonym Kept Five King Novels a Mystery". In: *The Bangor Daily News*, 1.
- Strachan, Don (Mar. 1981). "Soft Cover". In: *The Los Angeles Times*, 8.
- Straub, Peter (1984). "Meeting Stevie". In: *Fear Itself: The Horror Fiction of Stephen King*. Ed. by Tim Underwood and Chuck Miller. New American Library.
- Strengell, Heidi (2005). *Dissecting Stephen King: from the Gothic to literary naturalism*. University of Wisconsin Press.
- Thomases, Martha and John Robert Tebbel (Jan. 1981). "Interview with Stephen King". In: *High Times Magazine*.
- Tirvengadam, Vina (1996). "Linguistic Fingerprints and Literary Fraud". In: *Digital Studies/Le Champ Numérique* 2 (1). [10.16995/dscn.187](https://doi.org/10.16995/dscn.187).
- Tuzzi, Arjuna and Michele Alberto Cortelazzo (2018). "It Takes Many Hands to Draw Elena Ferrante's Profile". In: *Drawing Elena Ferrante's Profile*. Ed. by Arjuna Tuzzi and Michele Alberto Cortelazzo, 9-30.
- Tyo, Jacob, Bhuwan Dhingra, and Zachary C. Lipton (2022). "On the State of the Art in Authorship Attribution and Authorship Verification". In: *arXiv preprint*. [10.48550/arXiv.2209.06869](https://arxiv.org/abs/10.48550/arXiv.2209.06869).
- Underwood, Tim and Chuck Miller, eds. (1989). *Bare Bones: Conversations on Terror with Stephen King*. Warner Books.
- Van Cranenburgh, Andreas and Erik Ketzan (2021). "Stylometric Literariness Classification: the Case of Stephen King". In: *Proceedings of LaTeCH-CLfL 2021*. <https://aclanthology.org/2021.latechclfl-1.21.pdf> (visited on 11/29/2023).
- Williams, Nick B (Nov. 1984). "Thinner". In: *The Los Angeles Times Book Review*, 11.
- Winter, Douglas E. (Feb. 1985). "Stephen King, Peter Straub, and the Quest for the Talisman". In: *Twilight Zone Magazine*.

A. Appendix

A.1 Bachman attributions

Title	Rank	Harris	King	Koontz	Straub
<i>The Long Walk</i>	1	0.077	0.484	0.156	0.283
	2	0.164	0.282	0.241	0.313
	3	0.289	0.156	0.311	0.244
	4	0.471	0.077	0.292	0.160
<i>Roadwork</i>	1	0.125	0.400	0.171	0.304
	2	0.196	0.297	0.228	0.279
	3	0.284	0.194	0.288	0.235
	4	0.395	0.109	0.314	0.182
<i>Rage</i>	1	0.092	0.433	0.130	0.345
	2	0.174	0.308	0.221	0.298
	3	0.288	0.175	0.320	0.216
	4	0.446	0.084	0.329	0.141
<i>Blaze</i>	1	0.117	0.452	0.161	0.269
	2	0.202	0.282	0.226	0.289
	3	0.296	0.171	0.285	0.248
	4	0.384	0.095	0.328	0.193
<i>The Running Man</i>	1	0.132	0.368	0.211	0.289
	2	0.202	0.288	0.237	0.274
	3	0.284	0.208	0.268	0.241
	4	0.382	0.137	0.285	0.196
<i>Thinner</i>	1	0.087	0.452	0.177	0.283
	2	0.167	0.289	0.249	0.295
	3	0.282	0.172	0.299	0.247
	4	0.464	0.087	0.274	0.175
<i>The Regulators</i>	1	0.088	0.495	0.169	0.249
	2	0.176	0.271	0.252	0.301
	3	0.289	0.154	0.299	0.257
	4	0.447	0.081	0.280	0.192

Table 4: Bachman's book titles, ranks, and the proportions of 1,000-token segments from each book that were predicted to be written by each author in the corpus.

Title	Rank	Harris	King	Koontz	Straub
<i>The Long Walk</i>	1	0.010	0.733	0.052	0.205
	2	0.077	0.207	0.199	0.518
	3	0.326	0.052	0.408	0.213
	4	0.587	0.008	0.341	0.064
<i>Roadwork</i>	1	0.073	0.519	0.092	0.317
	2	0.180	0.275	0.168	0.377
	3	0.352	0.150	0.284	0.214
	4	0.395	0.056	0.456	0.092
<i>Rage</i>	1	0.036	0.561	0.045	0.358
	2	0.131	0.317	0.134	0.418
	3	0.367	0.098	0.367	0.168
	4	0.466	0.023	0.455	0.056
<i>Blaze</i>	1	0.033	0.702	0.056	0.209
	2	0.175	0.207	0.174	0.443
	3	0.402	0.068	0.303	0.226
	4	0.389	0.022	0.467	0.121
<i>The Running Man</i>	1	0.079	0.475	0.152	0.294
	2	0.203	0.247	0.209	0.341
	3	0.343	0.165	0.266	0.226
	4	0.376	0.113	0.374	0.138
<i>Thinner</i>	1	0.024	0.667	0.086	0.223
	2	0.102	0.234	0.223	0.441
	3	0.306	0.079	0.373	0.242
	4	0.568	0.020	0.318	0.093
<i>The Regulators</i>	1	0.008	0.818	0.049	0.125
	2	0.087	0.141	0.255	0.517
	3	0.300	0.034	0.396	0.270
	4	0.605	0.007	0.300	0.087

Table 5: Bachman's book titles, ranks, and the proportions of 10,000-token segments from each book that were predicted to be written by each author in the corpus.

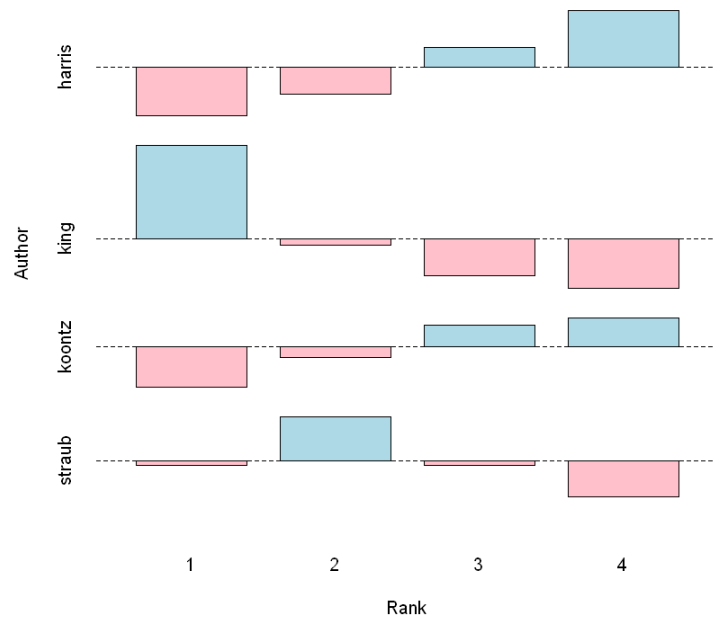
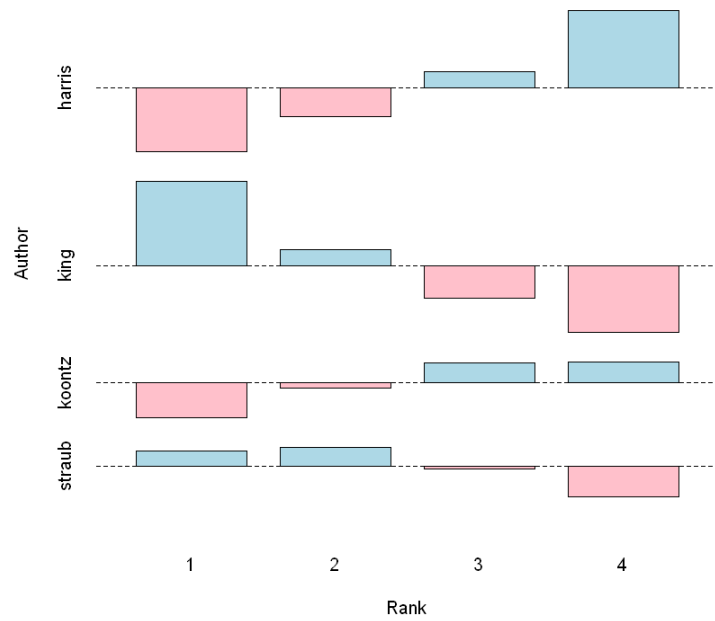


Figure 3: Association plots showing the direction and significance of correlation between the number of times a candidate author received a similarity ranking (1, 2, 3, or 4) to Bachman segments and candidate author in 1,000- (left) and 10,000- (right) token segments.

A.2 Pop-Culture References

A.2.1 Bachman

List of pop-culture references (brands, celebrities, products, fictional characters, movies, TV shows, etc.) extracted from *Rage* (1979), *The Long Walk* (1979), *Roadwork* (1981), and *Thinner* (1984).

A&S Tires, A. Gordon Pym, AP wire, Abdul Allhazred, Adreizi Brothers, Agatha Christie, Ahab, Albert Einstein, Alfie, Alligators All Around, American Express, Amoco, Amos 'n Andy, Amway, Anacin, Anaïs, And Justice for All, Andy Devine, Annie Oakley, Apple, Arco, Arlene Dahl, Art Linkletter, Aureomycin, Avis, B&O, BMW, Bach, Bally, Band-Aid, Banjo Rag, Banker's Life Insurance, Barbie, Bausch & Lomb, Be-Bop, Be-bop-a-lula, she's my baby, Beach Boys, Beatles, Beechcraft, Ben Alexander, Bermuda, Bertrand Russell, Beverly Hill-billies, Big Mac, Bill Cullen, Black Jack gum, Blackglama, Bob Hope, Bobby Sherman, Bombardier Skidoo, Bonneville, Brain from Planet Arous, Brian Wilson, Briggs & Stratton, Broderick Crawford, Bruce Springsteen, Bud, Budweiser, Buick, Burger King, Buttercup, Cadillac, Caesar's Palace, Calvin Klein, Camel, Campbell, Canada Mints, Captain Midnight, Captain Queeg, Chancellor-Brinkley, Chargers, Charles Manson, Charmin, Chatty Cathy, Cheez-Doodles, Chesterfield, Chevrolet, Chevy, Chevy Impala, Chevy Nova, Chipwich, Chivas, Chris-Craft, Chrysler Imperial, Chryster, Chuck Berry, Cimarron, Clint Eastwood, Coca-Cola, Coke, Colt, Colt Woodsman, Con-Tact, Cracker Jack, Curly, D-Con, Dairy Freez, Dallas Cowgirls, Dan Fortune, Dannon, Darvocet, Darvon, Datsun, David Cassidy, David Janssen, Delco, Delta 88, Denny's, Detroit Redwings, Detroit Tigers, Dial M for Murder, Dialing for Dollars, Diamond International, Dick Cavett, Dingo, Dior, Dirty Harry, Disney, Disney World, Dodge, Dodge Custom Cab, Dolby, Don Rickles, Donald Westlake, Dorito, Dorothy Sayers, Dos Passos, Dotto, Dr. Caligari, Dr. Scholl, Dragnet, Dungeons and Dragons, Duz, Dylan, Eames, Eberhard Faber, Econoline, Edith Head, Electrolux, Ellery Queen, Elton John, Elvis, Empirin, Ernest Hemingway, Exorcist, Exxon, F Troop, Facing the Lions, Family Feud, Fantasia, Farmer Brown, Fat Sammy's, Father Brown, Ferrari, Field and Stream, Firebird, Flair Fineliners, Flatt and Scruggs, Fountainbleau, Ford, Ford Pinto, Formica, Forrest Tucker, Francis Gary Powers, Frederick's of Hollywood, Frisbee, GI Joe, GM, Garfield, Garry Moore, Gary Davis, Ghidra, Gilligan's Island, Gisele MacKenzie, Glade Pine Fresh, Godzilla, Good-bye Yellow Brick Road, Goodwill, Gravy Train, Great Northern, Great Western, Green Door, Greenbriar Boys, Greyhound, Griff, Gucci, Gulf, Guy Lombardo, Guy Madison, H. Rider Haggard, HBO, Hal March, Hamburger Helper, Hammond Innes, Henry Glassman, Henry James, Henry Youngman, Herman Wouk, Hertz, Hesse, Hexlite, Hey, Mr. Sun, Highway Patrol, Hogan's Heroes, Holiday Inn, Home Box Office, Honda, Horchow, Hot Stuff, Howard Cosell, Howdy Doody, Hula Hoops, Humphrey Bogart, I Dream of Jeannie, I Love Lucy, Incredible Shrinking Man, J. C. Penney's, J. Press, J. W. Dant, J.C. Whitney, J.J. Cale, Jack Barry, Jack Benny Program, Jack Narz, Jack and Jill, James Bond, James Cain, Jaws, JayCee, Jell-O, Jerry Jeff Walker, Jimmy Cagney, Jimmy Stewart, Jingles, Jock Mahoney, Joe Friday, John Agar, John Carradine, John Chancellor, John Cheever, John Cougar Mellancamp, John D. MacDonald, John Travolta, Joker's Wild, Jordache, Judy Blume, Julia Child, KLH, Kalishnikov, Karen Carpenter, Kent, Kewpie, King Kong, Kitty Carlisle, Kleenex, Kluge, Kodachrome, Krazy Glue, Kurt Vonnegut, LTD, Lacoste, Larry, Lawrence Belch, Lee Strasberg, Lenny Bruce, Let

It Be, Let it Bleed, Let's Make a Deal, Levi's, Lionel Richie, Lipton, Lobstermen, Long Island Dragway, Loony Tunes, Lorne Greene, Louis Lamour, Mace, Malboros, Mammoth Mart, Manhattan Transfer, Mantovani, Marcus Welby, Marlboro, Marlboro Light, Marty Milner, Maurice Sendak, Mauser, Max Von Sydow, Maxwell House, McDonald, McDonald's, Mercedes, Mercury, Merv Griffin, Mets, Michael Jackson, Michelangelo, Mick Jagger, Mickey Mouse, Midnight Rambler, Milky Way, Miller, Miller Lite, Miracle Chopper, Moe, Molotov, Monkey Man, Monkey Trial, Monocle, Monopoly, Monte Hall, Morns, Mothra, Motorola, Mr. Bojangles, Mr. Hyde Jekyll, Mr. Rogers, Mr. T., Muenster cheese, Mustangs, Myron Floren, NFL, National Enquirer, National Geographic, National Lampoon, New Paltz, New York Times, Nike, Niques, Nite Owl, Nivea, Norman Bates, Norman Vincent Peale, Num-Zit, O. Henry, OK Corral, Olds, Oldsmobile, Ninety-Eight, Olivetti, Olivia Newton John, Omar the Tentmaker, Oshkosh, Outdoor Life, Pall Mall, Panthers, Pat Benatar, Paul Harvey, Paul Stuart, Pavlov, Peanuts comic strips, Peggy Sue, Penthouse, Pepsi, Pepsi-Cola, Pepto Bismol, Pequod, Perrier, Perry Mason, Peter Rabbit, Philco, Phillies Cheroot, Phillips, Pig Pen, Piper Cub, Plymouth, Polaroid, Ponderosa golf, Pontiac, Porsche, Psycho, Q-tip, Quoddy, RCA, REA express, Radio Shack, Ramada Inn, Range Rider, Raquel Welch, Ravi Shankar, Reader's Digest, Red Sox, Richard Petty, Richard Stark, Richard Widmark, Rin Tin Tin, Ring-Dings, Rinso, Ripley's Believe It Or Not, Ritz, Ritz crackers, Robert Redford, Rodan, Roloids, Rolex, Rollerdrome, Rolling Stones, Rolls-Royce, Ronald McDonald, S&H Green Stamps, SOI, Saab, Sadie Hawkins, Saf-T-Glass, Salvation Army, Samsonite, Sara Lee cheesecake, Saran Wrap, Saville Row, Schooner, Schwinn, Scripto, Sergeant Preston, Sears, Seven Flags Over Georgia, Seven-Up, Shell, Sheraton, Sherman tank, Shop and Save, Shop 'n Save, Shop 'n' Save, Slaughterhouse Five, Slurpies, Slurpy, Smith & Wesson, Sony, Soo Line, Soupy Sales, Souther Comfort, Southern Pacific, Space Command, Spencer Tracy, Spider John Koemer, Star Trek, Starsky and Hutch, Sterno, Stetsons, Sting, Stop and Shop, Stranger in Paradise, Subaru, Sunoco, Sylvester Stallone, T. S. Eliot, TRS-80, Tarr Brothers, Technicolor, Tensor, Tensor study lamps, Texaco, The \$64,000 Question, The Ballad of John and Yoko, The Day the Earth Stood Still, The Gift of the Magi, The Gong Show, The Guardian, The Manchester Guardian, The New Price is Right, The New York Review of Books, The Postman Always Rings Twice, This Is Your Life, This Savage Rapture, Thomas Carlyle, Thousand Island dressing, Three Stooges, Three's Company, Thunderbird, Thus Spake Zarathustra, Tic Tac Dough, Time Magazine, Timex, Tinkertoys, To Tell the Truth, Tolkien, Tom Paxton, Tom Rush, Tom Wicker, Tony Curtis, Toyota, Toys Are Joys, Trans Am, Trifles, Trix, True Argosy, Tukkan the Terrible, Twenty One, Twinkies, Twinky, Universal Pictures, VW, Vantage 100, Victor Canning, Vince Lombardi, Volvo, Von Ronk, WGAN-TV, Waldenbooks, Waldorf-Astoria, Walkman, Wall Street Journal, Walt Disney, Walter Cronkite, Warner Anderson, Warner Brothers, Washex, We Gotta Get It On Again, Weatherby, Wet-Nap, What's My Line, Where are They Now, Whoppers, Wild Bill Hickok, Wilt Chamberlain, Winslow Homer, Wizard of Oz, Woody Woodpecker, Wranglers, Wyatt Earp, Yankees, Yorick, You can't always get what you want, Your Hit Parade, Your Show of Shows, Zenith, Zippo, lazy Susan.

A.2.2 Harris

References extracted from *Black Sunday* (1975), *Red Dragon* (1981), and *The Silence of the Lambs* (1988):

280ZX, A Nurse to Marry, AK-47, Abdel Awad, Ain't Misbehavin, Air Force C-141, Aldo Ray, Alka-Seltzer, American Aermotor, American Express, Amex, Antoine's, Avon, Baby Ruth candy bar, Baeder Chemical, Band-Aid, Barry Manilow, Bartlett's Familiar, Baseball Joe, Batard-Montrachet, Beaujolais, Beaver Cleaver, Beechcraft, Begin the Beguine, Bell Atlantic, Beretta, Betty Skelton, Big Mac, Black Mountain Rag, Blondie, Bloomingdale, Boeing, Bolex Super Eight camera, Bonwit Teller, Britches, Bronco, Buffalo Bill, Bufferin, Buick, Bulldog .44 Special, C. S. Forrester, Camaros, Canoe after-shave, Canoe beer, Captain Video, Cardinals, Cash for Your Trash, Celotex, Cessna, Charles James, Chateaubriand, Checker, Chemical Mace, Cher Bono, Chevrolet, Cinzano, Citroën, Clorox, Coke, Cole Porter, Colt, Coromandel screen, Corsair, DC motor, Danny Kaye, Datafax, Deborah Harry, Decca, Delta, Demerol, Disney, Doc Watson, Dos Equis beer, Duccio, Duke Keomuka, Edith Piaf, e.e. cummings, El Diario-La Prensa, Elvis, Emily Dickinson, Erythromycin, Evelyn Waugh, Evyan, Ezio Pinza, F-4 Phantom, Fats Waller, Federal Express, Ferragamo, Flaying of Marsyas, Fleet, Flicket, Ford, Fotomat, Fox lock, Franklin Mint locomotives, GM, GOODYEAR DOUBLE EAGLE, Galatoire's, Garfinkel, Georgia Power Company, Glaser Safety Slugs, Glenn Gould, Greyhound, Grumman Gulfstream, Géricault, H. Allen Smith, Howard Hughes, Huckins, Huey, J. Edgar Hoover, Jack Daniel's, Jane Austen, Jell-O, Jimmy Hoffa, Johnny Carson, Joy of Cooking, Kaiser, Katyusha rocket, Kevlar, Kewpie doll, Kiss, Kleenex, Kodak D-76 developer, Kool-Aid, Kotex, L. L. Bean, Land Rover, Lean Cuisine, Lee Harvey Oswald, Levis, Lewis Carroll, LifeSaver, Lincoln Versailles, Listerine, Litton Policefax, Llama automatic pistol, Lomotil, Lord & Taylor, Lucite, Lutece, Lycra, Lysol, L'Air du Temps, Mace, Madonna, Mag Na Port, Magic Marker, Magnum, Man Mountain Dean, Marcus Aurelius, Mark Five gas mask, Mary Janes, Max Shulman, Melvin Purvis, Mets, Miami Dolphins, Miami Herald, Mike Hailwood, Moe, Monteleone Hotel, Montrachet, Morocco Mole, Mounds candy bar, Movietone News, Mr. Hide, NBS Sports Spectacular, NFL, National Football Conference, National Geographic, Nero, New York Times, Nikon, Nomex, Norman Vincent Peale, Novocaine, Odor-Eaters, Orkin, Over the Sea to Skye, Packard, Pan Am, ParkRite, Perelman, Perrier, Personality Plus, Peter Jennings, Phantom F-4, Phone-Mate, Picasso, Pinto, Pittsburgh Steelers, Plimsoll, Plutos, Plymouth, Polaroid CU-5, Port-O-San, Prince Andrew, Quonset, R. L. Polk and Company, Rand McNally, Reebok, Remington 870, Remy Martin, Reynolds 5130, Rice Stadium, Rinso white, Rinso bright, Ritalin, Rolodex, Romeos, Rose Marie Reid, Rubik's Cube, Rybovich, Saks, Sam Browne belt, Sancerre, Sapporo beer, Satellite Monroe, Schmeisser, Sears Best, Secret Squirrel, Sedan de Ville, Segovia, Servco Supreme, Shea Stadium, Sheetrock, Sikorsky, Cinderella, Skycrane, Smith & Wesson, Smithsonian's National Museum of Natural History, Smokey the Bear, Southeastern Bell, Southern Bell, Sperry-Rand, Sports Illustrated, Startron, Stevie Wonder, Sting-Eez, Studebaker, Styrofoam, Super Bowl, Superdome, Tanqueray, Tater Tots, Teflon, Telex, The Knifemakers Guild, The Look of Love, The Young and the Restless, Thorazine, Threave, Times, Titian, Toto, Trans-Am, Trumpy, Tulane Stadium, Tulane's Green Wave, Twinkie, UNICEF, United Coal, VanSleek Farfoon, Vanderbilt, Velcro, Vicks VapoRub, Visa, Vogue, Voice Privacy system, Volkswagen, Vonnegut, W. W. Greener, WPIK-TV, Walkman, Washington Post, Washington Redskins, Weight Watchers, Western Union, Whiskey River, Wile E. Coyote, Windex, Winston, Wolf's Ears, World Cup soccer, Wratten, Xerox, Yoo-Hoo, ZPG-1, Zamfir, Master of the Pan Flute, Zodiac, the Bargain Center, the Chicago Tribune, the China Mail, the Court of Two Sisters, the Daily News, the Fairmont Hotel, the Florida State League, the Gold-

berg Variations, the Harmon Trophy, the Holiday Inn, the International Herald-Tribune, the Intra Bank, the Los Angeles Times, the Marriott Hotel, the Monteleone Hotel, the National Broadcasting System, the National Gallery, the Navy's Blue Angels, the New Orleans Saints, the New York City Aquarium, the New York Post, the Oilers, the Ramlet el Baida, the Reader's Digest, the Royal Orleans Hotel, the Sugar Bowl Classic, the Super Bowl, the Tigers of Louisiana State University, the Times, the Top of the Mart, the World Series, the Yellow Pages.

A.2.3 King

References extracted from *Firestarter* (1980), *Cujo* (1981), and *Pet Sematary* (1983):

A & P, AFC Championship, AMC Matador, Ace bandage, Adidas, Adolph's Meat Tenderizer, Agway Market, Albany Airlines, Alfalfa, Ali MacGraw, Alka-Seltzer, All My Children, Allagash, Allegheny, AlpoBurgers, American Casket Company, American Express, Amex, Amoco, Amway, Anacin, Andrea Doria, Andy Warhol, Antonioni, Arco gas, Ariel Sharon, Armstrong ceiling, Arte Johnson, As the World Turns, AstroTurf, Atari, Atlanta Braves, Atlantic, Audrey Rose, Auld Lang Syne, Avis, Avon, B.J. and the Bear, Bally, Bang caps, BankAmericard, Barbie, Bass, Bat-Cycle, Bearcat scanner, Beatrix Potter, Becton-Dickson syringe, Bell helmet, Ben-Gay, Bentley, Bermuda onion, Bermudas, Bespin Warrior, Bette Davis, Beulah Land, Big Mac, Big Red Machine, Bijou, Bill Blass, Bird's Eye orange juice, Biscayne, Biz, Black Beauty, Black Label, Bloomingdale, Blue Cross-Blue Shield, Blue Horse tablet, Bluto, Bob Hope, Bob Seger, Bob Stanley, Bomba the Jungle Boy, Bone-Phone, Boris Karloff, Boston Post, Botany 500, Boy Scouts, Braniff Airlines, Brillo, Brookings-Smith Mortuary, Brooks Brothers, Buddy Hackett, Bugs Bunny, Buick, Burger King, Busch, Butterball, Cadillac, Caesar, Caldor, Camaro, Camel, Camera Store, Campbell, Canada Dry, Candy Man, Carlos Castenada, Casco Bank and Trust, Charles Dickens, Checker, Cheerios, Chester, Chesterfield Kings, Chevette, Chevrolet, Chevy, Chicago Tribune, Chrysler, Chuggy-Chuggy-Choo-Choo, Cisco, Citgo, Claymore, Clearasil, Clint Eastwood, Clio, Coca-Cola, Cocoa Bears, Coke, Colt, Con-Tact paper, Count Chocula, Crawly-Gator, Crayola, Cream, Credence, Cremora, Cuisinart, D. Duck, Dairy Queen, Dale Carnegie, Danskin leotard, Dark Victory, Darth Vader, Darvon, Dave Garroway, Dave and Frank Blair, Decoster Egg Farms, Dee Dee Ramone, Del Monte, Delco battery, Delta, Diamond matches, Dilly Bar, Diners Club, Dingos, Dinty Moore, Disney, Disney World, Doctor Doolittle, Dodge, Douglas MacArthur, Dow Chemical, Downy fabric softener, Dr. Cyclops, Dr. Denton suit, Dr. Seuss, Duke Wayne, Dukes of Hazzard, Dumbo, Dwight Frye, Eastern Airlines, Eastern Bank, Eeyore, Egg McMuffin, Ellsworth American, Elmer Fudd, Elmer's Glue, Elvis Presley, Encyclopedia Britannica, Erica Jong, Ernie, Esso, Ever-Lock, Exxon, Family Fun Lanes, Fay Wray's, Festus, Flexible Flyer, Ford, Forest Lawn, Franco Harris, Frankenberry, Frankenstein, GEN. PATTON, Gaines Meal, Gelusil, Gene Autry, Gene Simmons, General Hospital, George Carlin, George Romero, George and Gracie, Georgia Charger whiskey, Gerrypack, Gilbey's gin, Gillette Foamy, Gilligan's Island, Girl Scouts, Goofy, Gordon R. Dickson, Gravy Train, Greedo, Greyhound, Grover, Gucci, Gunsmoke, Hamburger Helper, Han Solo, Hefty bag, Heinz, Herbert Tareyton, Hershey, Hertz, Hillerich & Bradsby, Hitachi TV, HoJos, Holiday Inn, Home Box Office, Honda Civic, Hoosier cabinet, Horseman, pass by, Hot One Hundred, Hotpoint, Howard Johnson's Motor Lodge, Hush Puppies, Hush Puppy, IBM, Igor, Immelmann, Indian motorcycle, Injun Joe, International Harvester, Isodil, It's

a Good Life, J & B whisky, J. B. Rhine, J. C. Whitney & Co., J. Fred Muggs, J. J. Cale, J. Walter Thompson, J.R., JERRY FALWELL, Jack Daniel's, Jacob Marley, Jaguar, James Bond, Jaundaflo, Jeep, Jefferson Airplane, Jerome Bixby, Jerry Garcia, Jim Beam, Jim Morrison, Joan Baez, Jockey shorts, Joe DiMaggio, Joe Green, John Hurt, Johnny Carson, Johns Hopkins, Johnson's No More Tears shampoo, Johnson's Wax, Karl Malden, Keds, Keebler, Kellogg's, Kelvinator, Ken, Kermit, King Kong, Kleenex, Kodachrome, Kodak, Kool-Aid, L. L. Bean, La-Z-Boy, Lark cigarette, Latex, Laugh-In, Lawnboy, Lawrence Welk, Lee Riders, Lengyll, Lester, Lestoil, Lincoln Continental, Little Black Sambo, Little Golden Books, Little House on the Prairie, Lord Buxton, Lou McNally, Loudon Wainwright, Love Me Tender, Love Story, Love of Life, Lovecraft, Lucky Charms, Luger, Luke Skywalker, Löwenbräu, MIKE WALLACE, Magic Kingdom, Magic Mountain, Magnum, Mammoth Mart, Marek stove, Marshal Dillon, MasterCard, Matchbox, Matt Dillon, Maurice Sendak, Max Factor, McCheese, McDonald, Menachem Begin, Mercedes, Mercer Mayer, Michael Jackson, Micheloeb, Mickey Mantle, Mickey Mouse, Milky Way bar, Mille Bourne, Miller, Mixmaster, Mondavi, Monopoly, Mr. Coffee, Murray Leinster, Myer, NBC, Nabisco, Napoli's, Narnia, Necromancer, New Franklin Laundry, New York Mets, Nicklaus, Nipper, the RCA dog, Norman Rockwell, Northeast Bank, Noël Coward, Old MacDonald, Olympia, Omar the Tentmaker, Orasin, Orson Welles, Orville, Oscar, Oscar the Grouch, Oz the Great and Terrible, PATTI SMITH, PAUL HARVEY, Pall Mall, Pampers, Panasonic, Pancho, Parcheesi, Pearl Kineo, Pentel pen, Penthouse Forum, Pepsi-Cola, Peter Pan, Phil Donahue, Phone-Mate, Piggly Wiggly, Piggy, Pilot Razor Point, Pinto, Planet Mongo, Planet Quark, PlaySkool, Pledge, Plymouth, Polaroid, Ponch and John from CHiPS, Pooh, Popeye, Popov, Popsicle, Porsche, Puffer and Sons, Pyrex, Queen Victoria, Quell, Quonset, Raleigh, Ralston-Purina, Raspberry Zingers, Reader's Digest, Red Man, Red Rose bag, Red Sox, Redball Flyer, Revell airplane, Richard Dreyfuss, Richard Scarry, Ritz-Carlton, Roadrunner, Robbie the Robot, Robert A. Heinlein, Robert Gordon, Robert Parker, Rockaway Beach, Rolling Stone, Rolls-Royce, Rolodex, Roman Meal bread, Ronald McDonald, Rube Goldberg, Run Through the Jungle, SMERSH, Sam Cunningham, Sammy's Pizza, Sara Lee, Saran Wrap, Saturday Night Live, Scarlett, Schlitz, Schwinn, Scrabble, Scrooge, Sea World, Search for Tomorrow, Sears, Seeborg, Seiko, Sesame Street, Shakey's, Shakin' Stevens, Shedd's Peanut Butter, Sherlock Holmes, Shop 'n Save, Shuffle Off to Buffalo, Silex hotplate, Slim Jims, Smith & Wesson, Smith Brothers' Wild Cherry, Smucker's, Snackin' Cakes, Snickers, Snoopy, Sonny Bono, Sony, Space Invaders, Speedaway sled, Spic 'n Span, Spiderman, Spode china, Springsteen, Star Blazers, Star Wars, Starsky & Hutch, Sterno, Steve Martin, Stonehenge, Studebaker, Styrofoam, Sugaree, Sunkist, Sunoco, Superwoman, Sweet 'n Low, T. S. Eliot, TWA flightbag, Tarzan, Texaco, The Bangor Daily News, The CBS Morning News, The Cat in the Hat, The Chicago Tribune, The Creature from the Black Lagoon, The Crosswits, The Deer Hunter, The Deering Ice Cream Parlor, The Doctors, The Doobie Brothers, The Drac Pack, The Grateful Dead, The Headless Horseman, The Hundred Acre Wood, The Jimmy Durante Hour, The Kingston Trio, The Little Rascals, The Man from Glad, The Maytag Repairman, The Mellow Tiger, The Monkey's Paw, The Muppet Show, The New England Patriots, The New York Times, The New York Yankees, The New Zoo Revue, The PTL Club, The Pulitzer Prize, The Ramones, The Red & White grocery store, The Reds, The Rolling Stones, The Rookies, The San Diego Padres, The Super Bowl, The Tastee Freeze, The Temptations, The Tigers, The Today show, The Toledo Blade, The Tonight Show, The United Van Lines, The Washington Post, The Weapon Shops of

Ishtar, The Whitehall Hotel, The Wind and the Willows, The Young and the Restless, Thermos, This Ole House, Thorazine, Thunderbird, Tide, Tiffany box, Tiger tank, Tigger, Time magazine, Tipperary, Titanic, Toad, Tom Rush, Tom Watson, Tonka bulldozer, Tonto, Toonerville Trolley, Top Job, Trace Optical, Trinitron, Tuborg, Tuinal, Tupperware, Turtle Wax, Twinkie, U.S. of Archie, UN Plaza, Underalls, Underwood, United Airlines, United Cerebral Palsy, Upjohn, Utica Club beer, Valium, Van Donen, Van Vogt, Vantage cigarette, Vega, Vermont Maid Syrup, Victor Jory, Visa, W. W. JACOBS, WACZ, WCSH, WOYO, Walter Mitty, Watson, Watson's Hardware, Weight Watchers, Wells Fargo truck, Wendigo, Wheel of Fortune, Where the Wild Things Are, White Lightning, White Line Fever, Wilbur, Wilkie Collins, Willy Loman, Willy Wonka's Great Glass Elevator, Winchester, Winnebago, Winnie the Poë, Winston Churchill, Wonder Bread, Woolco, Wurlitzer jukebox, Wyeth, Wyman, Xerox, Zayre, Zenith television, Zig-Zag paper, Zippo.

A.2.4 Koontz

References extracted from *The Eyes of Darkness* (1981), *Phantoms* (1983), and *Darkfall* (1984):

7-Eleven, ABC, Abominable Snowman, Alan Alda, Alan Jackson, Albert Einstein, Alice in Wonderland, Alien, Amelia Earhart, American Express, Andrew Wyeth, Ann Landers, Bagley, Baloney, Barry Fitzgerald, Barry Manilow, Batman, Batmobile, Beatles, Beethoven, Bell JetRanger, Benny Goodman, Bermuda Triangle, Bernaise, Big Mac, Biosan-4, Botticelli Madonna, box of Cheer, Bulova, Busby Berkeley, CBS, Cadillac Seville, Cartier, Celica, Cessna, Charles Dickens, Charles Manson, Cheerios, Chevrolet, Chevy, Chewbacca the Wookie, Chivas Regal, Clairol, Coke, Coleman gas, Coors, Copernicus, Culligan, Dear Abby, Dennis the Menace, Dickens, Disneyland, Donatella, Dr Pepper, Dr. Faustus, Dracula, E.T., Eleanor Rigby, Electronic Battleship, Elmore Leonard, Elvis, Eroica, Explorer, Follett, Ford, Formica, Forsythe, Francis Bacon, Frank Sinatra, Frankenstein, Frosty the Snowman, Fudge Fantasies, Garth Brooks, General Electric, George Alexander, George Bernard Shaw, George Plimpton, Godzilla, Goodwill Industries, gouda, Gore, Graveyard, Groucho Marx, Gucci, Hallmark, Hank Thomas, Harley, Heckler & Koch, Honda, Hostess Twinkies, Howdy Doody, Irish Spring, Jack the Ripper, Jacqueline Bisset, Jalape 241, James Bond, Jasper Johns, Jeep, Joel Bandiri Presents, Judge Crater, K-Mart, Kleenex, Kraft Swiss cheese, Lalique, Land Rover, Lazarus, Levolor, Lexus, Life-Savers, Listerine, Lovecraft, Lyndon Johnson, M-1 semiautomatic, Ma Bell, MacLean, Magnum, Mario's Pizza, Marquis de Sade, Mary Celeste, McDonald's, Memorex, Mennen's Skin Conditioner, Mercedes, Mercedes-Benz, Mickey Mouse, Millionaire's Row, Mother Teresa, Mumm, Mussolini, NBC, Nash Rambler, New York Philharmonic, Noah, Norman Rockwell, O.J. Simpson, Pepsi, Plexiglas, Pontiac Trans Am, Pop Tarts, Pulitzer, Purina Cat Chow, Queen Anne, R. L. Stine, Rancho Circle, Raquel Welch, Remington, Remy Martin, Robert Redford, Rolex, Rolls-Royce, Rubik's Cube, Rudolph the Red-Nosed Reindeer, San Francisco Chronicle, Saran Wrap, Scott Baio, Sears, Seiko, Sharon Tate, Sheraton, Sherpa, Sidney Poitier, Silver Bells, Skylane RG, Smith & Wesson, Spam, St. Francis of Assisi, Star Wars, Superman, The Book of Job, The Cessna Turbo, The Mad Hatter, The New York Times, Thomas Mann, Time, Timex, Tiny Taylor, Tolstoy, Tom Dooley, Tonka Toys, Toyota, Tums, Vaseline, Vince Foster, Walt Disney, Walter Raleigh, Whiffle Ball, Wild Turkey, the Associated Press, the Bermuda Triangle, the New York Times, the Plaza Hotel, the Saturday Evening Post, the Twilight Zone, the Wall

Street Journal.

A.2.5 Straub

References extracted from *Ghost Story* (1979), *Shadowland* (1980), and *Floating Dragon* (1982):

A Praise of His Lady, ABC, Abraham Lincoln, Adidas, Adolf Eichmann, Agatha Christie, Alan Alda, Alan Ladd, Albert de Salvo, Amanda Cross, American Express, Andre Previn, Andrew Wyeth, Anne Bancroft, Anthony Powell, Aramis cologne, Archer Hotel, Archie Goodwin, Aretha Franklin, Arnold Palmer, Art Carney, Art Deco, Arthur Fonzarelli, Arthur Schlesinger, Audi, Audie Murphy, Audubon, BMW, Baldwin, Bambi, Bass Weejuns, Beatle, Ben Jonson, Ben Sidran, Benny Goodman, Betamax, Big Mac, Bill Perkins, Bill Terry, Bloomingdale, Blue Mountain beans, Bluto, Bo Diddley, Bobby Hackett, Bokhara rug, Bose, Bowie knife, Brooks Brothers, Brothers Grimm, Bruno, Bruno Hauptmann, Buck Rogers, Bud, Budweiser, Buick, Burberry, Burger King, Burl Ives, Butch Cassidy, CBS, Cadillac, Camaro, Campanella, Campari, Carrie, Cary Grant, Chaplin, Charles Addams, Charlie Antolini, Charlie Farrell, Charlie's Angels, Chaucer, Chevrolet, Chevy, Chiquita Banana, Christopher Isherwood, Cinderella, Claire Bloom, Clark Gable, Claude Rains, Coke, Constance Talmadge, Coors, Cornelius Agrippa, Cortez, Corvette, County Fair, Cuisinart, Currier and Ives, Cynara, D. H. Lawrence, Daimler, Daniel Boone, Dar Duryea, Death Star, Desi Arnaz, Diet Pepsi, Dingo, Disney, Dodge, Dolly Parton, Dom Perignon, Don Juan, Dorothy Sayers, Dr Faustus, Dr Pepper, Dracula, E. B. White, Eames chair, Edgar Allen Poe, Edvard Munch, Eliphaz Levi, Elizabeth Jane Howard, Ella Fitzgerald, Elvis Presley, Emma Bovary, Ernest Dowson, Ernest Hemingway, Ernie Kovacs, Eugene Palette, Exxon, F. SCOTT FITZGERALD, F. W. Dixon, Falada, Flexible Flyers, Florence Nightingale, Fludd, Ford, Foreign Intrigue, Four Roses, Frank Sinatra, Frankenstein, Fred Astaire, Freud, Frosty the Snowman, Gary Cooper, Gatsby, Gene Shalit, General Pershing, George Shearing, Glenn Miller, Golden Chicken, Goodwill, Grace Bumbry, Great Expectations, Gremlin, H. P. Lovecraft, Halston dress, Hank Williams, Hansel, Harpo Marx, Harry Carey, Jr., Harry Truman, Henry James, Herman Wouk, Holden Caulfield, Houdini, Huckleberry Finn, Humphrey Bogart, Hush Puppies, I Love Lucy, IBM, Ichabod Crane, Ilie Nastase, Isobel Archer, Ivan the Terrible, Ivy Compton-Burnett, J. D. Salinger, JOHN BARRYMORE, Jack Nicholson, Jackie Gleason, Jaguar, James Bond, James Cagney, James Dean, James Fenimore Cooper, James Stewart, Jameson, Jane Pauley, Janet Gaynor, Jann Wenner, Jean Harlow, Jean Rhys, Jell-O, Jim Beam, Jim Reeves, Jimmy Durante, Jimmy Nervo, Joan Crawford, Joe McCarthy, John D. MacDonald, John Dean, John Denver, John Ford, John Gilbert, John Held, John Kennedy, John Scarne, John Updike, Johnnie Ray, Johnnie Walker, Johnny Carson, Johnny Unitas, Johnny Walker Black, Josh Randall, Joyce Carol Oates, Katharine Hepburn, Keds, Kitaj, Kiwanis, Kleenex, Knute Rockne, La Belle Dame Sans Merci, La Grande Illusion, Labatt, Laura Ashley, Le Baron, Lewis and Clarke, Liane D'Eve, Lincoln, Links Golf, Lion's Club, Little Red Riding Hood, Lonnie Donegan, Loretta Lynn, Louise Brooks, Lully, MG, Madame Blavatsky, Madame Bovary, Mae West, Magic Fingers, Maudie Scott, Manson family, Margaret Drabble, Margaux, Marilyn Monroe, Mark Hopkins, Marlboro, Mars bar, Mary Astor, Mary Miles Minter, Mary Pickford, Matchbox, Mather, Mazda, McDonald's, Mello Yello, Melvyn Douglas, Mercedes, Miami Herald, Mickey Mouse, Midnight Tango, Miss Marple, Monopoly, Monsieur Verdoux, Morgan, Mount

Rushmore, NBC, Nansen, Napa Valley Chardonnay, Nathaniel Hawthorne, National Broadcasting Company, Nero Wolfe, Night of the Living Dead, Noble Sissle, Norma Shearer, Norman Rockwell, Nosferatu, O. Henry, Old Parr, Oliver Hardy, Oreos, Orson Welles, Ouija board, Packard car, Pampers, Pandora's Box, Pansy Osmond, Paul Horning, Paul Scott, Paul Stuart, Pepsi, Pepsi-Cola, Pernod, Perrault, Peter Lorre, Peter Pan, Peugeot, Picasso, Piggly Wiggly, Pissarro, Pitch 'n Putt golf, Playboy, Plaza Hotel, Pocket Books, Polka Dots and Moonbeams, Ponce de Leon, Princeton University Press, Pulitzer-prize, Puvis de Chavannes, R. D. Jameson, R. P. Blackmur, Randolph Scott, Rasputin, Raymond Chandler, Remington, Remy Martin, Renoir, Rex Bell, Rex Stout, Rex, the Wonder Horse, Rialto theater, Richard Barthelmess, Richard Speck, Rip Van Winkle, Robert Burns, Robert Ferguson, Robert Frost, Robert Redford, Robert Reed, Rolling Stone, Rosa Forte, Rose Room, Rotary, Sam Shepard, Scott Fitzgerald, Sears, Sergeant York, Shaker chair, Sheldon Leonard, Shiraz carpet, Sir Walter Scott, Smith and Wesson, Snickers, Snoopy, Snowy Breasted Pearl, Song for a Sucker Like You, Sony, Spencer Tracy, Spiderman, Starsky and Hutch, State Farm, Stephen Crane, Stetson, Steve McQueen, Steve Miller, Stilton, Stolichnaya, Studebaker, Stychen Tyme, Sunoco, Surfin' Bird, Sweet Sue, Ted Koppel, Teddy Knox, Texas Ranger, The Alamo, The Archer Hotel, The Brady Bunch, The Cheshire Cat, The Elks, The Everly Brothers, The Far Side of Paradise, The Garrick Club, The Hands of Dr. Orlac, The Hardy Boys, The Honeymooners, The House of the Seven Gables, The Invisible Man, The Jaycees, The John Birch Society, The Kansas City Times, The Key of Solomon, The Legend of Sleepy Hollow, The Little White Cloud That Cried, The Mad Hatter, The Making of a Surgeon, The Marble Faun, The Million Dollar Roundtable, The Narrative of A. Gordon Pym, The National Rifle Association, The Odyssey, The Phil Donahue Show, The Red Badge of Courage, The Rhetoric of Irony, The Scarlet Letter, The Statler Hilton, The Three Musketeers, The VFW, The Wizard of Oz, There's a Small Hotel, Thomas Mann, Tiffany lamp, Tom Brokaw, Tom Seaveris, Tommy Flanagan, Tony Archer, Tottel's Miscellany, Toyota, Treasure Island, Tupperware, Tweedledum and Tweedledee, UPS, Uri Geller, Valium, Van Helsing, Vandyke, Vanity Fair, Vanny Chard, Vergil, Village Pump restaurant, Vilma Banky, Vincent Price, Virginia Woolf, Volvo, W. C. Fields, W. H. AUDEN, Wabash Cannonball, Waldenbooks, Waldorf-Astoria, Walter Cronkite, Wayne Booth, When The Red, Red Robin Goes Bob, Bob, Bobbin Along, William Bendix, William Powell, Willie Mays, Willie Nelson, Wilton, Wimbledon, Winnebago, Woodward and Bernstein, Wyatt Earp, Xerox, YPSL, Yoda, Young Brothers department store, Zoot Sims.

A.2.6 Pop-Culture References Excluded from Bachman

Pop-Culture Reference	Reason for Exclusion
Larry	Character in Stephen King's <i>The Stand</i>
Buick	Character in Stephen King's <i>From a Buick 8</i>
Jingles	Character in Stephen King's <i>The Green Mile</i>
Sears	Character in Peter Straub's <i>Ghost Story</i>
Bud	Character in Peter Straub's <i>Shadowland</i>
Dylan	Character in Dean Koontz' <i>By the Light of the Moon</i>
Mace	Character in Dean Koontz' <i>Warlock</i>
Campbell	Character in Dean Koontz' <i>The Husband</i>

Table 6: Pop-culture references found in Bachman's books that were excluded from analysis with their reason for exclusion.

A.3 Statistics

Group	N	W	p
Bachman	700	662,439	<0.001
Straub	1,200		
Bachman	700	1,011,395	<0.001
Koontz	2,000		
Bachman	700	963,633	<0.001
King	2,000		
Bachman	700	274,381	<0.001
Harris	500		

Table 7: Results of Wilcoxon rank-sums one-sided tests comparing pop-culture reference counts extracted from randomly-sampled 10,000-token segments in Bachman, Straub, Koontz, King, and Harris books. Tests compared pop-culture reference counts in Bachman and Straub segments, Bachman and Koontz segments, Bachman and King segments, and Bachman and Harris segments.

Group	N	W	p
King	2,000	1,560,287	<0.001
Straub	1,200		
King	2,000	2,222,412	0.004
Koontz	600		
King	2,000	639,999	<0.001
Harris	500		

Table 8: Results of Wilcoxon rank-sums one-sided tests comparing pop-culture reference counts extracted from randomly-sampled 10,000-token segments in King, Straub, Koontz, and Harris books. Tests compared popular reference counts in King and Straub segments, King and Koontz segments, and King and Harris segments.

Authorship Attribution of Stephen King's 'Bachman' Books

Author	Book Title	Mean	Me- dian	Mini- mum	Maxi- mum	Standard De- viation
Bachman	<i>Blaze</i>	5.95	4	0	15	3.63
	<i>Rage</i>	20.12	18	7	48	8.32
	<i>Roadwork</i>	23.96	21.5	8	51	11.18
	<i>The Long Walk</i>	1.81	2	0	4	1.24
	<i>The Regulators</i>	3.29	3	0	19	3.15
	<i>The Running Man</i>	1.52	1	0	6	1.87
	<i>Thinner</i>	16.83	13	5	43	8.70
Harris	<i>Black Sunday</i>	1.41	1	0	4	1.22
	<i>Hannibal</i>	2.48	2	0	9	1.94
	<i>Hannibal Rising</i>	0.77	0.5	0	3	0.89
	<i>Red Dragon</i>	1.48	1	0	7	1.91
	<i>The Silence of the Lambs</i>	3.24	3	0	10	2.49
King	<i>Bag of Bones</i>	5.34	4	0	23	5.38
	<i>Carrie</i>	2.79	2	0	8	2.10
	<i>Cell</i>	3.10	3	0	8	2.22
	<i>Colorado Kid</i>	4.46	5	1	8	1.83
	<i>Desperation</i>	2.12	2	0	8	1.81
	<i>Dreamcatcher</i>	3.39	2	0	13	2.89
	<i>From a Buick 8</i>	3.23	3	0	11	2.51
	<i>Insomnia</i>	3.79	3	0	11	2.54
	<i>Lisey's Story</i>	4.69	4	0	14	3.68
	<i>Rose Madder</i>	2.90	2	0	12	3.11
	<i>Salem's Lot</i>	3.12	3	0	6	1.84
	<i>Song of Susannah</i>	2.67	2	0	9	2.67
	<i>The Dark Tower</i>	1.98	2	0	7	1.69
	<i>The Dead Zone</i>	4.42	4	0	13	2.80
	<i>The Girl Who Loved Tom Gordon</i>	13.51	14	6	25	4.19
	<i>The Green Mile</i>	1.23	1	0	5	1.41
	<i>The Shining</i>	2.63	2	0	10	2.41
	<i>The Stand</i>	5.42	4	0	29	6.28
	<i>Wizard and Glass</i>	0.74	0	0	10	1.62
	<i>Wolves of the Calla</i>	1.39	1	0	9	2.11
Koontz	<i>After the Last Race</i>	2.97	2	0	13	3.46
	<i>Beastchild</i>	0.22	0	0	1	0.42
	<i>By the Light of the Moon</i>	3.49	2.5	0	12	3.20
	<i>Fear Nothing</i>	1.39	1	0	6	1.42

Continued on next page

Table 9 – Continued from previous page

Author	Book Title	Mean	Median	Minimum	Maximum	Standard Deviation
	<i>From the Corner of His Eye</i>	3.92	3	0	13	3.43
	<i>Hideaway</i>	5.47	5	1	17	3.42
	<i>Lightning</i>	7.92	5	2	29	7.02
	<i>Night Chills</i>	1.14	1	0	4	1.28
	<i>Phantoms</i>	1.30	1	0	11	2.01
	<i>Star Quest</i>	0.00	0	0	0	0.00
	<i>Strangers</i>	3.51	3	0	20	3.24
	<i>The Bad Place</i>	4.67	3	0	33	5.81
	<i>The Eyes of Darkness</i>	5.04	4	0	14	3.89
	<i>The Good Guy</i>	5.93	3	0	28	6.21
	<i>The Husband</i>	7.41	7	0	15	4.51
	<i>The Taking</i>	1.30	1	0	5	1.37
	<i>The Vision</i>	4.11	4	0	10	2.74
	<i>Warlock</i>	0.00	0	0	0	0.00
	<i>Whispers</i>	2.31	1	0	11	2.81
	<i>Winter Moon</i>	3.35	2	0	18	4.07
Straub	<i>Floating Dragon</i>	2.42	2	0	10	2.31
	<i>Ghost Story</i>	1.03	1	0	5	1.18
	<i>Hellfire Club</i>	2.41	1	0	14	3.42
	<i>If You Could See Me Now</i>	1.76	2	0	4	1.20
	<i>In The Night Room</i>	3.32	3	0	11	2.87
	<i>Julia</i>	0.40	0	0	3	0.98
	<i>Koko</i>	2.69	2	0	16	3.02
	<i>Lost Boy, Lost Girl</i>	1.47	1	0	4	1.42
	<i>Mr. X</i>	1.55	1	0	7	1.27
	<i>Mystery</i>	2.02	1	0	11	2.29
	<i>Shadowland</i>	0.83	0	0	4	1.04
	<i>The Throat</i>	2.75	2	0	13	2.84

Table 9: Mean, median, minimum, maximum, and standard deviations of pop-culture reference counts extracted from randomly sampled 10,000-token segments from the Bachman books and books in the distractor corpus.

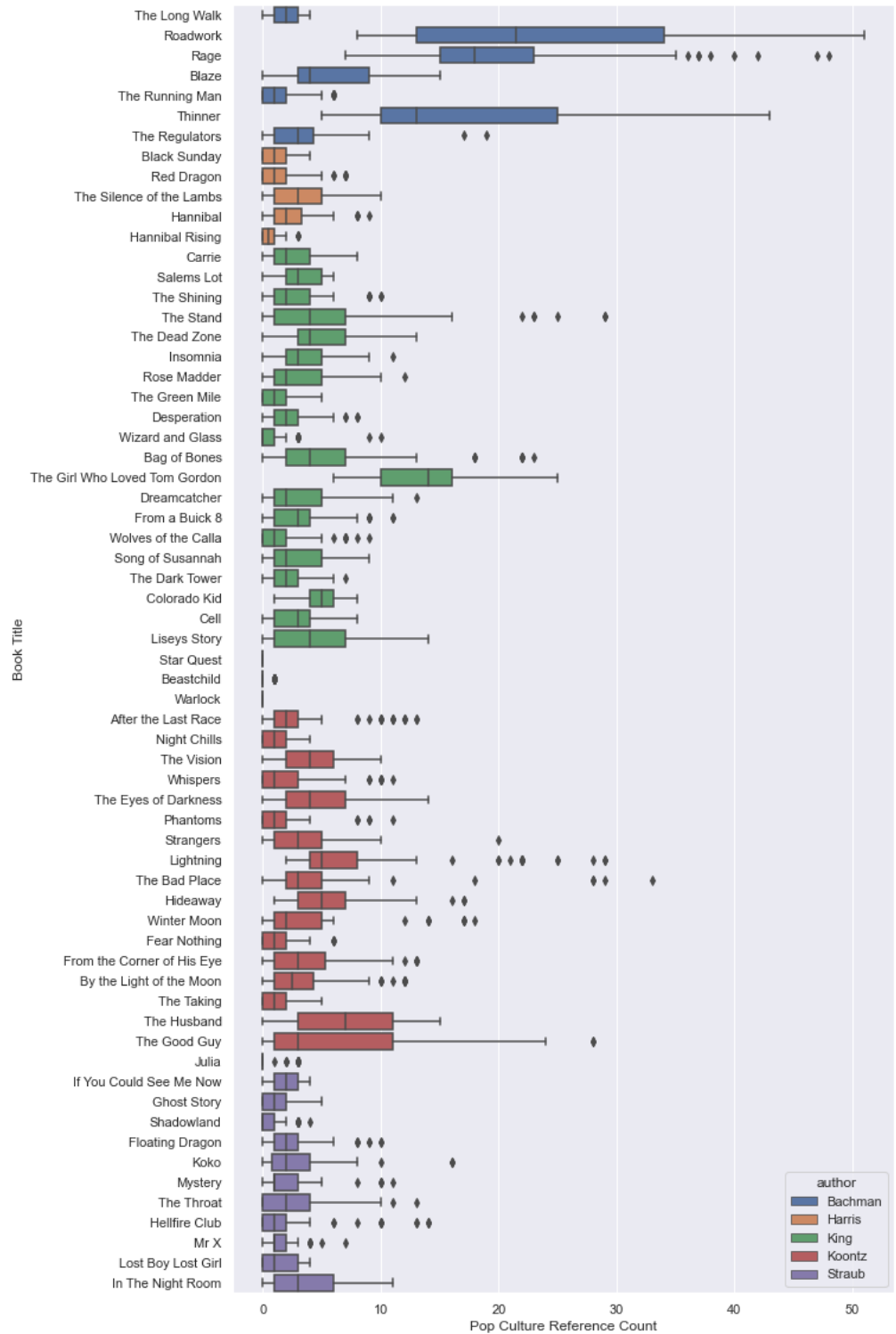


Figure 4: Boxplot visualizing pop-culture reference counts by book title. Counts were extracted from randomly sampled 10,000-token segments from the Bachman books and books in the distractor corpus.