

**MAN PAIN IN THE MAN BOOKER PRIZE:  
A QUANTITATIVE APPROACH  
TO CONTEMPORARY CANON FORMATION**

by

CAITLIN E. POWELL

SUBMITTED TO SCRIPPS COLLEGE IN PARTIAL FULFILLMENT  
OF THE DEGREE OF BACHELOR OF ARTS

PROFESSOR JACQUELINE WERNIMONT  
PROFESSOR WARREN LIU

APRIL 25, 2014

And what stood in their way? Their personalities and pasts, their ignorance and fear, timidity, squeamishness, lack of entitlement or experience or easy manners, then the tail end of a religious prohibition, their Englishness and class, and his history itself. Nothing much at all. He removed his hand and drew her to him and kissed her on the lips, with all the restraint he was capable of, holding back his tongue.

- IAN MCEWAN, *ON CHESIL BEACH*

## Table of Contents

Introduction .....	4
Methodology .....	15
Results and Discussion .....	24
<i>Author-based Metadata</i> .....	24
<i>Presentation-based Metadata</i> .....	34
<i>Formal Variables</i> .....	41
<i>Content-based Metadata</i> .....	46
<i>Structural Variables</i> .....	64
Conclusion .....	80
Works Cited .....	82

## Introduction

### *Canonicity and Best Texts*

Much of our cultural interaction with both contemporary literature and literature from the past is defined by the impulse to label certain texts as the best texts available to be read, regardless of the legitimacy or the necessity of this impulse, and it is this impulse that underlies the process of literary canon formation. Ankhi Mukherjee writes that “Canonicity implies the formation of a corpus...the making up of a list of books requisite for a literary education, and the formation of an exclusive club, however painstakingly contested the rules of inclusion (and exclusion) may be” (1029) – in other words, the formation of a literary canon is the assembling of a set of best texts, a set of texts that are deemed of higher intellectual quality and greater critical value than the texts excluded from the set. Indeed, when we think of a traditional literary canon, we think of a group of works that have been classified as worthy of reading, studying, and criticizing, over and over again by hundreds upon hundreds of people. A work’s inclusion or exclusion from this group determines whether or not that work is later read, disseminated, and discussed, and, in some cases, whether or not that work endures in the long term.

The question then remains as to what criteria may be used to determine a work’s inclusion or exclusion from a canon or a set of best texts, and John Guillory’s *Cultural Capital: The Problem of Literary Canon Formation* is particularly useful in attempting to answer it. As Guillory argues, the traditional approach to canon formation is to think of canonical works as representing the voices of a community and non-canonical works as representing excluded or minority voices. Guillory, however, asserts that canon formation is not about actively selecting groups to exclude and groups to represent based on a set of



values held by the selectors, or even about the content or reputation of individual works; rather, “Canonicity is not a property of the work itself, but of its transmission, its relation to other works in a collection of works -- the syllabus in its instructional locus, the school” (Guillory 55). Because literary studies have ceased to have much monetary or market value and because, in terms of simply ensuring literacy, they have been replaced by composition, the canon is functional primarily as a reproduction of a classed notion of value. Since books cannot actually be divided neatly into “worthy” and “utterly useless,” the canon is nothing but an imaginary list that is transmitted through institutions of power, like schools, and the instruments thereof, like syllabi. Schools and professors regulate access to literacy and make judgments of literary value – literary value is determined through reproduction of social ideas of value, not through active decisions to represent or fail to represent certain voices.

Following this line of argument, if I want to better understand how contemporary literature, specifically contemporary novels, are canonized, how do I go about examining which values are being reproduced and transmitted? Many canons – 19<sup>th</sup> century British literature, for example – are already well defined in list after list and syllabus after syllabus, but who judges what contemporary literature is worthy of being read and studied and what syllabus reproduces this list of worthy contemporary literature? Because I don’t have access to such an existing syllabus, in order to answer these questions, I need both a metric for sorting novels that are categorized as the best novels and novels that are not and a way to analyze the sorted best novels en masse for commonalities, trends, and traits that make those works specifically and uniquely high quality in their temporal context. To address the first of these points – the need for a way to filter works that have

been labeled the best from those that have not – this project will begin by investigating the ways in which the Booker Prize may serve as an appropriate, if highly and perhaps overly selective, tool for this purpose.

### *The Booker Prize*

The Booker Prize – or, as it is now known, the Man Booker Prize – is a literary prize for the “the best novel of the year written by a citizen of the United Kingdom, the Commonwealth or the Republic of Ireland” (“The Man Booker Prizes”). A recent rule change allows for consideration of novels written in English by authors who are citizens of any country in the world, including the United States, so long as the novel was published in the UK, but this project will not mirror this change. The winning novel is selected from a shortlist of five to six novels by a panel of judges, including literary critics, writers, publishers, and other academics, and is announced every October at an extravagant ceremony at London’s Guildhall. The winner receives £50,000 and the authors of the novels that do not win but make the shortlist receive £2,500; the purpose of the prize, as currently stated by the prize’s website, is “to encourage the widest possible readership for the best in literary fiction” (“The Man Booker Prizes”). Authors who win, as well as authors whose novels make the shortlist, see dramatic increases in the sales of the winning or shortlisted novel, as well as increases in sales of their back catalogues and increases in the popularity and perceived legitimacy of their future works.

Of the 100,000 titles published in the last twenty years that were eligible for the Booker Prize, only 2,500 were submitted by publishers for consideration (Todd 9), and only approximately 250 titles have been shortlisted in the entirety of the Booker Prize’s

history; clearly, for the purpose of examining a set of contemporary texts that are generally classified as the best texts available, the novels nominated for the Booker Prize cannot function as the entire set of best English-language texts worth examining. Rather, these novels function as a sample of a much larger population of novels. Based on the definition of canonicity offered previously, the Booker Prize functions as one of many potential creators of contemporary canon, and the novels that win or are shortlisted constitute a sample of the best contemporary novels; in Guillory's terms, we may consider the list of Booker Prize nominees a continually growing syllabus, a syllabus that is reproducing the judgments and values of the Booker Prize Foundation as an institution and the judges that make the selections. By creating this syllabus, it is creating a list of books that should be read, and thus is regulating access to a certain kind of contemporary literacy – books that are not nominated are obviously still available for purchase, but because they are excluded from the syllabus, they are less likely to be widely read, widely recognized, and widely adopted as of quality by the literate, classed public who reads contemporary literature in the first place.

If the corpus of Booker Prize shortlisters and winners is a representative cross-section of the kind of novel that is deemed of highest quality and highest value, then examining the traits common to Booker Prize nominees is a way to access traits common to the set of contemporary best novels and thus potential criteria for inclusion in that set. That is, if I use the Booker Prize as a way to select the best novels, I can examine the traits of Booker Prize nominees in order to determine what constitutes and qualifies as a best novel. Additionally, as the Booker Prize, prior to the recent rule change, dealt with English and Commonwealth novels only, the set of best texts it creates and determines is

pleasingly continuous with the traditional English-language literary canon, which is very much grounded in British texts, rather than American texts or non-Western texts.

Of course, being nominated for a literary prize or even winning a literary prize does not guarantee canonicity. History demonstrates over and over that texts considered valuable in their time are often forgotten or discredited ten or fifty or a hundred years later; texts that we now consider crucial members of literary canons were not necessarily respected or even read during the time period in which they were written. However, I have no such temporal distance from a contemporary canon and thus no way to access the list of texts written in this time period that will be valued fifty years from now; texts that are nominated for the Booker Prize may not be considered important or literarily valuable in the future. However, the group of novels nominated for the Booker Prize does give me a way to begin constructing a potential model of a contemporary best text.

#### *Novel as Subjective Consumer Product*

The novels included in the set of best texts as established by the Booker Prize are not necessarily best texts in any objective, transcendent sense. Their status as best texts is entirely determined by the tastes and political choices of the panel of judges selecting them and, as a result of their selection, their increased readership, distribution, and intellectual profile. And because there is no such thing as a group of texts that would undeniably be deemed the best texts ever written by any conceivable combination of people assessing them, it is precisely this idea of achieving canonicity through increased readership and prominence that makes the Booker Prize a particularly effective prize in pre-selecting a sample of best contemporary texts.

Why, though, would the Booker Prize be any more effective at this selection process than any other literary prize, of which there are many? The answer lies in money and sales. As Richard Todd writes in *Consuming Fictions: The Booker Prize and Fiction in Britain Today*, “contemporary literary canon-formation is subject to powerful, rapidly changing market forces affecting and influencing the consumer” (9). That is, contemporary fiction is, in many ways, a consumer product – the books that are read are the books to which readers have access, and readers have access to the books that are thoroughly publicized, advertised, marketed, and distributed. If canon is in part commercial, as it has always been – before art was bought by the general public, it was commissioned by patrons – it matters very much to canon formation which books are bought, sold, and perceived as significant and, before the Booker Prize, literary prizes did not affect this. Before the Booker Prize, “there were prizes to be won by the serious literary novelist...but their significance was not noticed by the majority of the reading public, nor were they promoted as being of interest to consumers of contemporary fiction” (Todd 55); that is, these prizes did nothing to really increase either the readership or the perceived legitimacy of the novels they recognized. While they may have determined a set of best texts, they did little to create a pervasive set of best texts that was representative of both the kind of novel contemporarily perceived as being of high quality and the kind of novel that would be influential on the novels written after it.

However, in 1964, just as the idea for the Booker Prize was just coming into being, Tom Maschler, a member of a panel of speakers at a National Book League meeting concerning the future and purpose of literary prizes, argued that the purpose of a literary prize was “firstly, a means of showing appreciation for a writer; secondly, a

means to aid writers financially; and thirdly, a way to enhance a writer's reputation and consequently increase sales" (Norris 21). When the Booker Prize finally emerged in 1969, it was centered on money, prestige, and increased sales and prominence of the novels it recognized; the monetary reward associated with winning was £5,000, a considerable sum at the time and more than was associated with any other literary prize. This significant prize money was meant to have the effect of raising the profile and media coverage of the books and authors who won the Booker Prize, rendering those books more important and relevant to the British literary scene than other books, even other books that won other literary prizes (Norris 22). This is because the stakes are higher for a literary prize with a larger monetary reward – it matters more to win a prize whose reward constitutes a large amount of a good with high use value, and thus the books that win this prize matter more. This creates a cycle of importance – as the books that win begin to matter more, it becomes even more meaningful to win than it was previously, and so on in a circle.

Thus, the Booker Prize, because it not only creates a set of best texts but disseminates them in a way that ensures that they are read, helps us to identify a sample of canonical contemporary texts in a unique way. If the Booker Prize did not affect book sales, it would not put novels, previously unheard of or otherwise, on academic, popular, and critical maps with the same efficiency and potency. Ultimately, because the corpus of nominees selected by the Booker Prize judges becomes widely understood and widely known as representative of the best contemporary novels, I can use the sample created to extrapolate to the population and argue that the literary traits common to the Booker Prize

nominees are also literary traits common to the best contemporary novels as a broader group.

*The Booker Prize and Cultural Literary Preference*

Does the Booker Prize actually function as an accurate predictor of the kinds of books that become valuable, though? Do the books that are shortlisted for the Booker Prize actually represent the kinds of books perceived as literarily valuable? Milada Frankova says yes, writing that, “the Booker has proved to be a reliable and sensitive barometer of the times in terms of the ever-changing literary aesthetic and cultural values. The Booker Prize winning novels as well as those on the short-list can therefore conveniently fill temporary fiction curricula and consequently build new canons” (53). That is, she believes that the Booker Prize functions as an accurate metric to filter books that represent the artistic values of the culture and society of the time. A literary canon should, ideally, shed light onto the culture that created it, and the Booker Prize, as a determinant of a contemporary canon, does this.

Merrit Moseley, who has written and reported on the Booker Prize selection process for many years, argues that “while celebrating the best novel in English by eligible novelists is the official aim, stirring up conflict and outrage is important, too. Perennial arguments have bubbled over whether the book should be readable or literary (implying difficult and unpopular); over the role of foreign authors versus the homegrown; and over whether judges should scout for new, unknown talent or reward veterans” (“Margins” 429). Supporting Frankova’s claim that the Booker Prize serves as a sensitive measure for literary aesthetic, this indicates that the novel eventually picked as

the winner, as well as the novels eventually picked for the shortlist, will reflect a certain cultural literary bias – it will either reflect consistencies and commonalities among the novels or ways in which those commonalities evolve and change.

Mosley gives several illustrative examples of how the character of the shortlist reveals constant or evolving trends in the set of best texts recognized by the Booker. For example, if, in a period of a decade, shortlist after shortlist is filled with the same kinds of novels by the same kinds of authors, that says something important about the kind of novel being recognized as among the best novels being written. If shortlists repeatedly consist of novels by middle-aged white men, or novels set in 1950s London, or novels featuring the coming-of-age processes of young English men, this reveals a distinctly different kind of best text than if shortlists repeatedly consist of novels written by female authors of color, or novels set in Africa, or novels featuring the struggles of a mother-daughter relationship. If shortlists feature the same authors over and over, this says something different than if shortlists regularly recognize previously unknown authors. Consistent trends reveal underlying cultural biases regarding the kind of novel perceived as a best text.

However, the Booker Prize shortlists from year to year can also reveal the evolution of the kind of novel perceived as a best text. For example, in 2009, the shortlist was suddenly and unexpectedly dominated by historical fiction and the winner, Hilary Mantel's *Wolf Hall*, was set in the Tudor Period and featured Thomas Cromwell as its protagonist; this was surprising because at this point, historical fiction as a genre was previously considered un-literary and unlikely to win any significant literary prizes, due to its failure to address contemporary life, struggles, and concerns (Moseley, "Margins,"



430). Additionally, in 2006, the shortlist comprised of brand-new novelists that no one had ever heard of – the winning author, Kiran Desai, was fresh out of a creative writing program, and many heavyweight novelists like Martin Amis were longlisted but culled in the creation of the shortlist (Moseley, “Tidy and Untidy,” 289). In 2007, the shortlisted books were unusually short, concise, and “tidy” – Ian McEwan’s *On Chesil Beach* was barely more than a novella, clocking in at 200 pages with unusually few words on each page (Moseley, “Tidy and Untidy,” 299). In all three of these cases, the shortlist both reflects shifting attitudes and gives values to texts that might not have otherwise been valued; the sudden recognition of a work of historical fiction or a particularly short work both demonstrates a change in the kind of novel considered valuable and worthy of the set of best texts and influences the kind of novels accepted as valuable in the future, because the novels recognized by the Booker Prize may change how novels are written after they are recognized.

### *Objective of the Current Project*

Clearly, a single year’s shortlist can reveal both a continuation of a previously existing trend or a shift indicating the beginnings of a new trend. However, what happens if I examine the trends, both consistent and shifting, across multiple shortlists of many consecutive years? What patterns emerge if I look at the 254 Booker Prize winners and shortlisters through 2012 in their totality, and interrogate this larger sample of novels in order to create a model of what constitutes a contemporary best novel?

Of course, analyzing approximately over two hundred novels is a daunting task if one approaches it with traditional methodology – in *Consuming Fictions*, Richard Todd,

expresses this very concern, writing that in order to examine the group of novels that comprise the Booker Prize winners and shortlisters, he “set about constructing [his] own much smaller corpus in the belief that it is better to say more about fewer than the other way around” (9). However, how is it possible to get a comprehensive, representative view of the consistent and shifting characteristics of these books, the characteristics that qualify them for both the prize and membership in the set of best texts in the first place, by only examining a few of the books? How is it possible to guarantee that the one or three or ten books selected were the books that exemplified traits common across the entire corpus, traits that pervaded a significant number of the texts at hand and traits that made those texts particularly and uniquely valuable in their contemporary setting? I believe that a better approach is to do what Todd appears to say that it is not possible to do – to say more about more – and that is what this project will attempt to do.

## Methodology

### *Traditional Scholarship*

Traditional literary scholarship focuses on the exemplary. A scholarly essay might focus on one text – one poem, one novel, or one short story, for example – or perhaps two or three, with the intent of comparing different texts by the same author or different texts from the same time period or different texts with a similar theme. As digital humanities scholar Franco Moretti puts it in his book *Graphs Maps Trees*, literary scholarship functions as a series of case studies (3); traditional scholars examine the exceptional, texts deemed unusual and distinct and extraordinary, and they look for the sections and sentences and words that are exceptional within those exceptional texts. While there is nothing wrong with this approach, and while it has immense merit if the goal is to examine the efficacy or staying power or potency of a particular work, it fails to address anything remotely resembling the totality of work produced of a certain kind (British novels of the 19<sup>th</sup> century, American poetry of the 20<sup>th</sup> century, Booker Prize shortlisters). As Moretti points out, “a canon of two hundred novels, for instance, sounds very large for nineteenth-century Britain (and *is* much larger than the current one), but is still less than one percent of the novels that were actually published” (4). The samples of texts deemed representative of a certain population of texts may not, in fact, be representative of that population at all – the two hundred novels scholars examine over and over from different angles when they examine 19<sup>th</sup> century British novels may not accurately reflect the overall character of the totality of the novels produced during that time period. In terms of the Booker Prize, the two or three or ten or twenty best known shortlisted or winning

novels might not accurately reflect the commonalities, trends, and overall character of the entire population of shortlisted and winning novels.

As Matthew Jockers comments in his book *Macroanalysis: Digital Methods and Literary History*, part of this largely idiographic approach is due to the methodology traditionally adopted by humanities disciplines, literary studies included. This methodology is one that relies on careful, detailed, focused investigation – literary studies “close reads” and plumbs a text for the nuances and subtleties of its construction, and as such gathers evidence in a way that leads necessarily to subjective interpretation. While some readings of a text may be unsubstantiated and unsupported, one text may lead to multiple profound, interesting, and equally valid interpretations, which may either contradict or complement one another. This methodology doesn’t leave room for examining large quantities of texts at once, and it “resist[s] or avoid[s] computational approaches to the study of literature” (4); that is, if the methodological focus is the subjective examination of the meaning made through unique combinations of words and sentences and themes and characters, there is no room left for quantitatively analyzing texts from a distance or for removing analysis from the realm of the uniqueness of the individual text. Close reading is a deep method of evidence-gathering, but it is not a broad one – a vast amount of evidence will go unexamined simply because each of us cannot close read every single text written in a ten year period, or a fifty year period, or a century.

*Distant Reading and Macroanalysis*

Moretti, also in *Graphs Maps Trees*, offers a model of literary analysis that addresses the historical absence of quantitative data analysis in the field, and Jockers, also in *Macroanalysis*, offers a comprehensive explanation of the value of this model. Moretti's model is called distant reading, and it is an attempt to reduce texts and render them purposefully abstract in order to achieve "fewer elements, hence a sharper sense of their overall interconnection. Shapes, relations, structures. Forms. Models" (1). Moretti "reads" many texts superficially, rather than one or two texts closely; he does not interpret the texts and instead gathers data from them, looking at, for example, the number of novels published each year in Britain in the 1700s and displaying those numbers visually ("the rise of the novel"). He distances himself from the content of the texts themselves and focuses instead on characteristics and features of the text that are apparent without reading them – rather than looking at how a word functions in a specific text, we might look at the frequency with which that word is used across one hundred texts and chart the change in that frequency over time.

Jockers compares distant reading to macroeconomics and close reading to microeconomics; where microeconomics focuses on individual businesses and consumers, macroeconomics focuses on the behavior and fluctuations of an entire economy, operating at a much higher level than microeconomics (24). Similarly, macroanalysis of literature operates on a broader, more abstracted level than microanalysis – examining an entire genre or an entire time period gives a sense of a fluctuating whole, rather than a sense of the intricate behaviors of an individual. As Jockers argues, this ability to examine a whole rather than an individual gives us distinct

advantages and allows us to answer questions that we would not be able to answer otherwise; for example, we are able to look at how texts and authors fit into broad historical contexts and patterns, we are able to look at literary production and growth over time, we are able to look at changes in word use and writing style and theme over time, and, most relevant to my particular project, we are able to examine “the tastes and preferences of the literary establishment” (27). Ultimately, macroanalysis (or distant reading – the two terms function extremely similarly) allows us to construct comprehensive models of certain kinds of texts; in the case of my project, it allows me to propose a model of the best contemporary novel and back that model up with evidence from a large body of texts.

#### *Analytical Variables and Kinds of Data*

In doing this kind of distant reading on Booker Prize winners and shortlisters, the first step is to determine exactly what kinds of abstracted variables might possibly be analyzed. Jockers lays out several categories of data to help with this question, the most useful of which, for my purposes, are metadata and theme.

Metadata is bibliographic or catalog data –that is, information concerning the surface-level characteristics of the body of work contained within a “comprehensive literary bibliography” (Jockers 36), rather than information on the particular books within those bibliographies. This kind of data, as Jockers puts it, is useful because “in the absence of full text, this bibliographic metadata can reveal useful information about literary trends” (36) – that is, when a full text is present, our inclination is to focus on individual authors, individual texts, and the rich and detailed information contained

within those individual texts. However, without a full text to examine, we're left with data concerning the author (gender, nationality, birthplace), or the novel's publication (location, year), or the novel's setting (urban, rural, in what country). For example, Jockers describes the process of developing a database of metadata concerning Irish American prose spanning 250 years, and discusses how examining categories like author gender, author birthplace, and rural vs. urban setting allowed him to trace the evolution of a set of texts on very broad axes in a way that examining specific chunks of text within specific works would not have.

Thematic data, on the other hand, deals with the content of a large volume of individual works – this can be anything from common plot structures to certain recurring types of characters to thematic threads (such as “building confidence,” “coming of age,” or “learning to deal with death”) that run through the corpus as a whole. We might track certain plot features, such as young men encountering mentors or setting out on journeys; we might examine how frequently certain characters, like an evil mother figure or a wise old man figure, appear across the texts in the corpus; we might, as Jockers suggests, examine keywords relating to a certain thematic concept running through the corpus. That is, if we want to examine the theme of “confidence” in a set of texts, we can identify words related to the idea of confidence and track their frequency and usage throughout the texts. Ultimately, this type of data analysis allows us to examine not only the qualities of the corpus as an entity in and of itself, as metadata enables us to do, but allows us to examine both the consistencies and the shifts of the contents of the corpus.

## *Approaches of the Current Project*

### Author-based Metadata

The first kind of data I will examine is author-based metadata, which is metadata concerning the characteristics of the group of authors who wrote the novels shortlisted for the Booker Prize; this kind of data is important because it reveals the kinds of voices included in the corpus and thus the viewpoint and perspective towards which the novels are skewed. In particular, I will look at the number of women on the shortlist, the number of authors who were nominated more than once, and the locations in which the authors were born; I will represent the percentage of the shortlist made up of women per year with a line graph, the number of repeat nominees with a bar graph, and the birth locations with a shaded map.

### Presentation-based Metadata

Presentation-based metadata includes data on how a novel is aesthetically packaged and presented; specifically, in this case, what its cover looks like in terms of color scheme used, image type used, and font style used, all three of which will be represented as bar graphs. This kind of data, while it does not necessarily reveal anything critical about the content of the novel, does reveal how the consumer of the novel is meant to perceive the novel. The cover of a physical book is seen before anything else – before the inner content is seen, before the words are read, perhaps even before the blurb or quotes on the back are read. Thus, the cover of the book primes the consumer and gives the consumer clues about what “type” of book is at hand.



For the purposes of this project, all data is taken from the covers featured with the books in the archive on the Booker Prize website, rather than from the covers of the novels' first editions; this is because the covers presented on the Booker Prize website represent how the Booker Prize Foundation is currently choosing to present the book on its own website, which might a better indication of the qualities of the book the Booker Prize Foundation wanted to emphasize, rather than the qualities a certain publishing house wanted to emphasize.

### Formal Variables

For this project, the formal variable I've chosen to analyze is average page length of the first editions of the novels per year, depicted as a line graph. This is not metadata, because it looks at qualities of the contents of the novels, but is also not a structural variable, because it is not examining the plot or thematic content of the novels; rather, it examines variations on an element of novelistic form. Novels of different lengths may have different qualities or genre associations and understanding how long the average Booker Prize shortlister is may help in capturing these qualities.

### Content-based Metadata

Content-based metadata deals with the content of the novels but does not go into thematic content – rather, in the case of this project, it looks at where the novels are set geographically, what kinds of settings the novels feature (rural, urban, etc.), what time period the novels are set in, and how many protagonists of different genders the novels feature. Geographic setting will be depicted as a shaded map; setting type, time period,

and number of protagonists will be represented in bar graphs; and percentage of the shortlisted novels containing at least one male protagonist per year will be represented in a line graph. This kind of data is important because it offers me an understanding of the framework in which the stories are operating – the kinds of characters acting out the story at hand and the environment and era in which they are doing so.

### Structural Variables

Structural variables give me an understanding of how the novels are structured thematically – that is, what kinds of stories are being told within the framework established by the content-based meta variables. In order to examine the plots and thematic content, I will look at the descriptive blurb that accompanies each novel in the Booker Prize online archive – this is a more useful strategy than simply examining the blurb on the back of a copy of the novel, as the blurbs on the Booker Prize website, as with the covers, are the blurbs chosen to accompany the novels in an explicitly Booker Prize-related context, rather than the blurbs chosen to accompany the novels by the novels' publishers. Examining these blurbs tells us how the Booker Prize Foundation has chosen to present the novels it selects, which in turn gives us more accurate information about what elements of the novels the Booker Prize Foundation has deemed important and noteworthy.

To analyze the blurbs, I identified all the adjectives in each, since adjectives, as opposed to nouns or verbs, are more likely to capture a novel's theme or atmosphere. Hundreds of blurbs could use the noun "man," for example – the telling part is whether the man is described as evil or haunted or sad or angry or British, etc. Once I collected all

the adjectives, I divided them up into eighteen categories, which I created based on the kinds of words I noticed appearing most frequently. I will represent these categories in a bar graph.

### Individual Texts

Finally, in order to illustrate what distant reading can illuminate or miss within individual texts, I will incorporate analysis of three separate novels within my analysis of the data collected on the entire corpus. So as to include novels by both men and women, novels from multiple decades, and novels that represent both winners and shortlisters, I will look at Margaret Atwood's *Cat's Eye* (shortlist, 1988), Kazuo Ishiguro's *The Remains of the Day* (winner, 1989), and Ian McEwan's *On Chesil Beach* (shortlist, 2007). All three of these authors are repeat nominees – Ishiguro has been nominated four times and Atwood and McEwan five times each – which means that their novels reliably exemplify qualities deemed particularly worthy or valuable, and closely analyzing their work alongside data that has been removed from individual texts will help me look at individual texts within a more informed framework.

## Results and Discussion

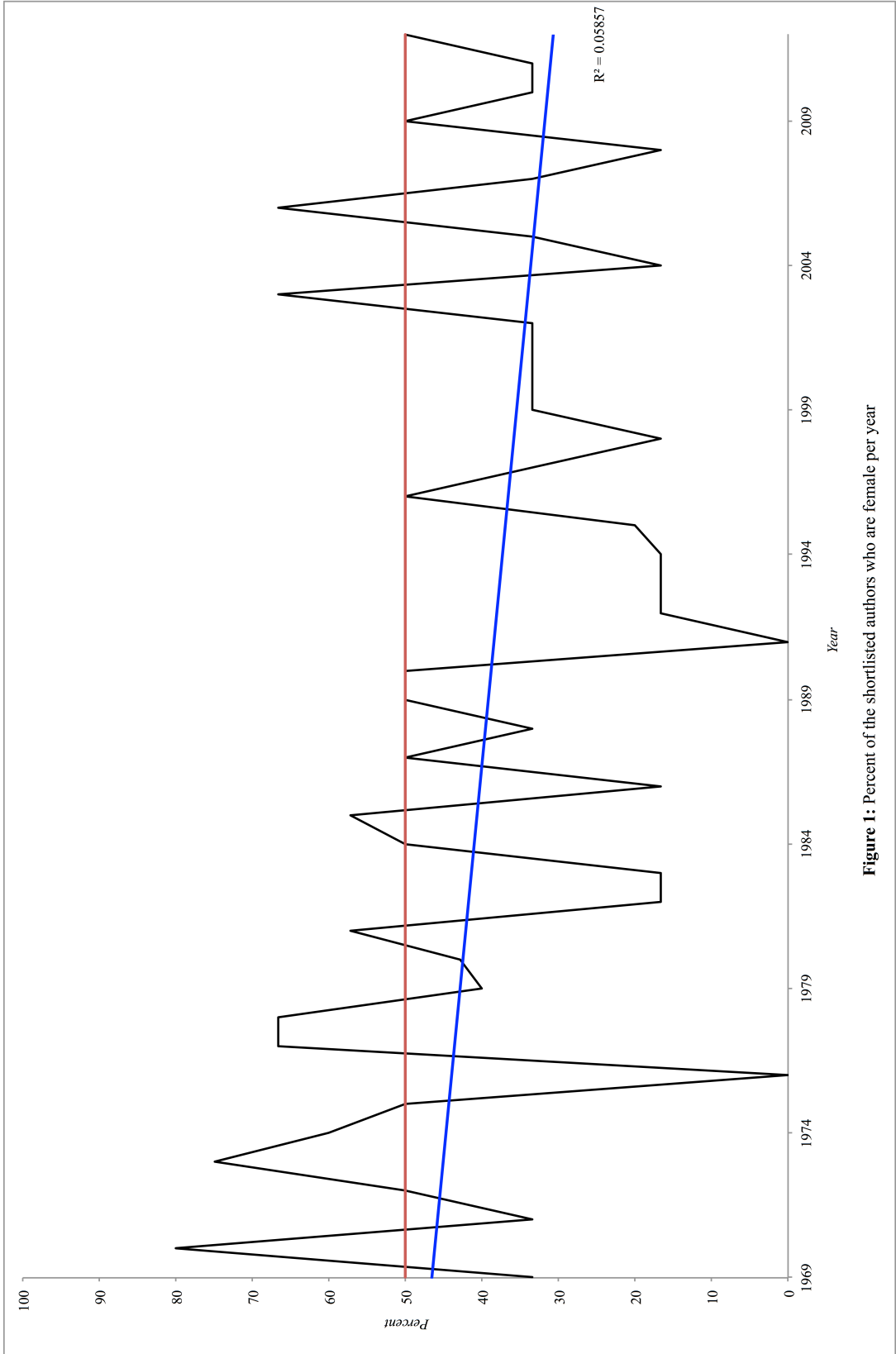
### *Author-based Metadata*

#### Female Authors on the Shortlist

Figure 1 demonstrates the prevalence of female authorship among the 254 novels in the corpus. The black line represents the percentage of shortlisted authors that were female in a given year; when the black line reaches higher than the red fifty percent line, the majority of the authors in that year were female, and when the black line fails to reach the red fifty percent line, the majority of the authors in that year were male. Figure 1 reveals that in most years, female authors make up less than fifty percent of the shortlisted authors. In two instances, shown by the black line touching the x-axis of the graph, the shortlist contained entirely men, and the shortlist has yet to include entirely women or, indeed, over eighty percent women. Figure 1 also shows massive variation in the percent of women from year to year – however, because the shortlist is small enough that the large visual difference between sixty percent women and eighty percent women is generally one author, the massive variation depicted on the graph is visually much more dramatic than it is when simply examining the list of authors.

Additionally, the trend line on this graph, depicted as a blue line, indicates that the number of women on the shortlist has decreased slightly with time; that is, since 1969, the percentage of the shortlist made up of women has been gradually getting smaller. This trend is not a strong one – the coefficient of determination, or  $R^2$  value, indicates that only 5.857% of the variation in the data is accounted for by a linear relationship between time and percentage female authors. However, in combination with the yearly low percentages of female authors, it is clear that the novels that win or are shortlisted for

the Booker Prize largely represent male voices, rather than female ones. The perspective on the world offered by this body of literature is a male-dominated one, and one that grows increasingly more masculine with time; it appears that rather than simply selecting the best novels of a given year written by citizens of the United Kingdom, the Commonwealth, or the Republic of Ireland, the Booker Prize is biased toward selecting the best novels of a given year written by male citizens of the United Kingdom, the Commonwealth, or the Republic of Ireland. Long before examining attributes of the novels themselves, it is clear that the idea of “best” is a gendered one.



**Figure 1:** Percent of the shortlisted authors who are female per year

## Repeat Nominees

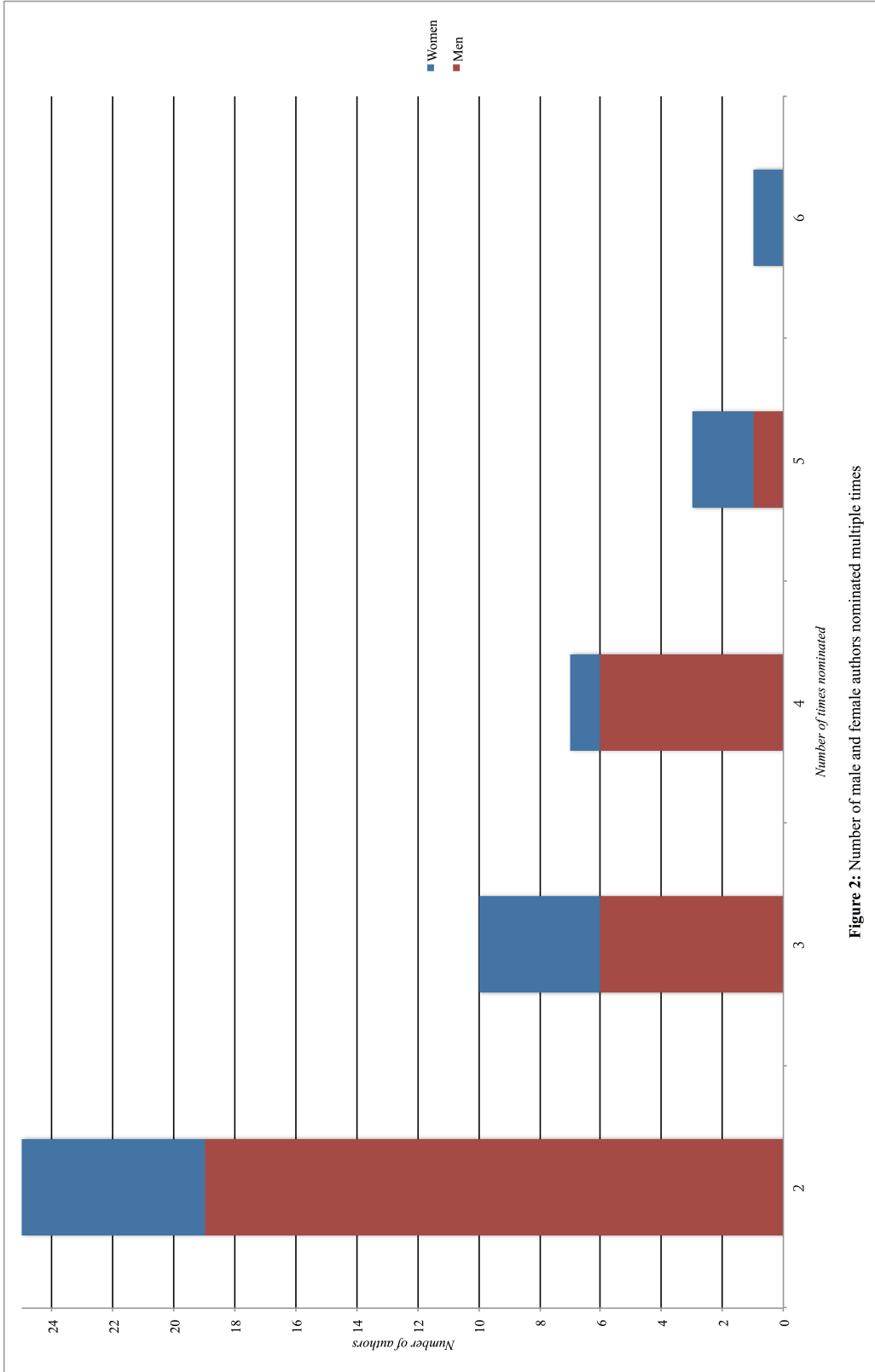
Despite the selective and prestigious nature of the Booker Prize, multiple authors, both men and women, have been nominated anywhere from two to six times – for the 254 novels that have been nominated since 1969, there are only 171 authors. Figure 2 demonstrates that twenty five authors have been nominated twice, ten have been nominated three times, seven have been nominated four times, three have been nominated five times, and one has been nominated six times, and also shows that in most cases, the striking majority of these repeat nominees are men. The only cases in which the majority of repeat nominees are women are the five time and six time nominee cases; Margaret Atwood and Beryl Bainbridge were nominated five times, along with Ian McEwan, and Iris Murdoch is the only author to have been nominated six times. Interestingly, Beryl Bainbridge’s most recent nomination was in 1998 and Iris Murdoch’s was in 1987, while Margaret Atwood’s was in 2003 and Ian McEwan’s was in 2007; this lends some support to the trend of decreasing recognition for female authors demonstrated in Figure 1, as all of the repeat nominations for Murdoch and Bainbridge occurred before 2000 and all of McEwan’s occurred after 1990.

The high number of male repeat nominees suggests a distinct, established group of male writers whose voices establish and permeate the corpus of nominated novels. While the Booker Prize recognizes new authorial voices, it also returns over and over to the same group of writers; a writer has a better chance of being nominated if they have been nominated before, if their work has already been recognized as the “best” of contemporary literature written in English. This large group of repeat nominees suggests a kind of cyclical quality of the Booker Prize – it reinforces its own selections in that the

more an author is nominated, the more worthy their work will be perceived and the more likely he or she is to be nominated again.

Additionally, the fact that female nominees are the majority only in the five time and six time nominee categories indicates an acceptance of the tokenized successful female writer, rather than acceptance of the less-famous talented female writer. A small number of extremely successful, well-known female writers can be nominated over and over and still not threaten the large number of male writers who are nominated two or three times – these few women who are nominated such an usually high number of times are exemplary and extraordinary, rather than typical, and their frequent nominations suggest both an acceptance of a small number of exemplary female writers and a resistance to a large number of two or three time female nominees.





**Figure 2:** Number of male and female authors nominated multiple times

### Author Birthplace

Figure 3 and its inset show where the authors nominated for the Booker Prize were born; if an area on the map is shaded in, at least one author was born in that area, and the darker the shading, the more authors born there. The restriction requiring that Booker Prize nominees be citizens of the United Kingdom, the Commonwealth, or the Republic of Ireland obviously limits the birthplace possibilities, but a distinct trend is evident nonetheless.

The Commonwealth of Nations includes 53 member states, most of which are former British territories or colonies; it was formed as the British Empire dissolved and many occupied states gained their independence. Today, Commonwealth citizens make up approximately one-third of the world population. However, of these 53 member states comprising one-third of the world's population, only thirteen are represented (the United Kingdom, Canada, Australia, New Zealand, Trinidad, Tanzania, South Africa, India, Malaysia, Nigeria, Sri Lanka, Pakistan, Bangladesh, and Ghana), despite all 53 being eligible birthplaces for nominated authors. Eight nations that are not part of the Commonwealth at all (the United States, Iran, Egypt, the West Indies, China, Japan, Germany, and Spain) are also on the list of birthplaces of nominated authors; the authors born in these nations simply became citizens of the United Kingdom later in life.

Ultimately, what this means is that only approximately 25% of the Commonwealth nations have produced even one author who was nominated for the Booker Prize. Additionally, of the 171 authors, approximately 66% of them were born in the United Kingdom (England, Scotland, Wales, and Northern Ireland) or the Republic of Ireland; of this 66%, approximately 28% were born in London alone (19% of the total

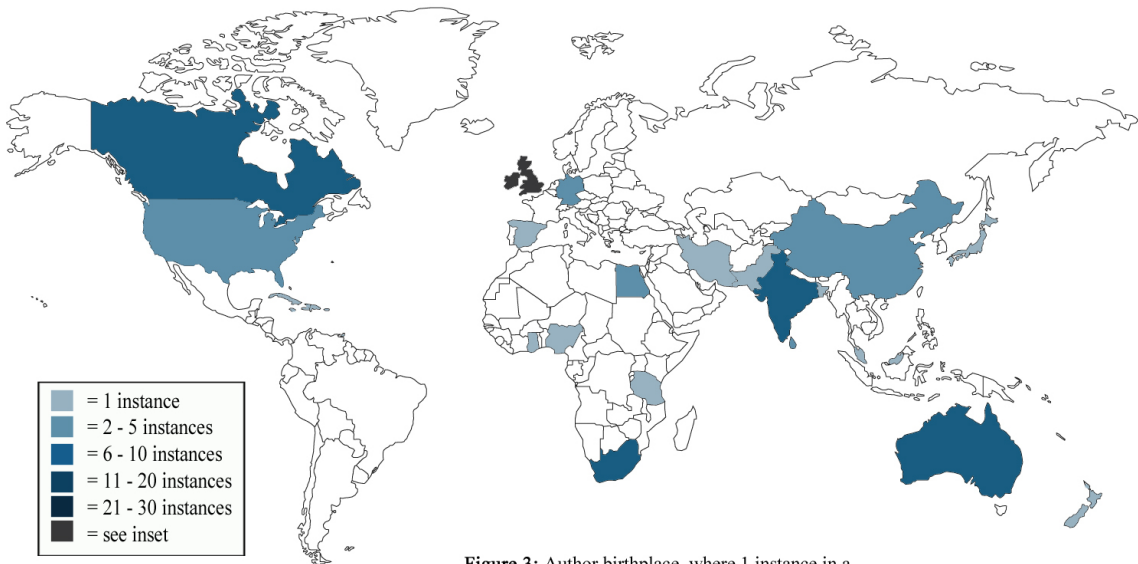
171 authors). Of the remaining 34% of authors not born in the UK or Ireland, approximately 6% were born in India, 5% in South Africa, and 4% each in Canada and Australia. This leaves only 15% of authors remaining, and this 15% makes up all the authors born in the rest of the Commonwealth and non-Commonwealth states.

Furthermore, the inset map of the British Isles in Figure 3 reveals even more trends. The darkest areas on the map are London and Northern England; the lightest areas are Wales and Cornwall, with Scotland and Southern England as the second lightest and Middle England and both Northern Ireland and the Republic of Ireland as the second darkest. Clearly, even as the United Kingdom is dramatically overrepresented among Commonwealth nations, certain areas of the United Kingdom are much better represented than others. London, the epicenter of British national identity, is especially overrepresented, not just among the authors born in the UK but among all 171 authors total. Similarly, locations less connected with Britishness and British nationalism, like Wales and Cornwall, which have traditionally been excluded from full and equal membership in British and/or English identity, are underrepresented and have produced relatively few nominated authors (five and two, respectively).

This geographical bias has serious implications regarding the voices and perspectives present in the novels nominated for the Booker Prize, both in terms of politics and in terms of identity. To begin with, the novels selected are dominated by a British perspective and, taking into account Figures 1 and 2, a British male perspective; the colonizers, rather than the colonized, are writing the books that are deemed the best literature produced in English by citizens of the UK, the Commonwealth, and Ireland. Even the requirement that the literature be written in English biases selection toward the

wealthy, English-speaking countries, and thus towards the wealthy, Western, colonizing perspective. Even setting aside the United Kingdom, of the other better-represented states, Canada, Australia, and New Zealand are all developed, Westernized, and relatively financially and politically secure; the perspectives of citizens of these nations are not very likely to be perspectives of oppressed or colonized groups, of marginalized people or people of color. They are likely to be white, middle- to upper-class perspectives.

Additionally, of the better-represented areas with a greater history of oppression and hardship at the hands of England and the British Empire – namely, South Africa, India, and the Republic of Ireland – these states are all infamously tied to British identity, British national crisis, and British imperial decline. Once again, the novels nominated for the Booker Prize are skewed towards authors whose origins are connected with Britishness, with nations and states famous for struggles that involved England rather than struggles entirely separate from England and its associated whiteness as a colonizing force.



**Figure 3:** Author birthplace, where 1 instance in a location represents one author born in that location



## *Presentation-based Metadata*

### Color Scheme

Figure 4 demonstrates how frequently certain color schemes appear as the dominant color scheme on covers of Booker Prize winners and nominees. The most common color schemes are sepias, darks, and earth tones, followed by primary colors, reds and blacks, photographs, deep solids, acid greens and yellows, and pastels; infrequently appearing color schemes include neon colors and rainbow designs, pure white covers or white covers with a single black line drawing, sunset colored covers, and printed covers (camouflage, a national flag, the page of a comic book, or a vintage-style illustration).

To clarify what the most frequent categories mean specifically, color schemes marked as “sepia” are any characterized by the aged-photograph sepia effect, in which drawings or photographs are printed with browned ink. “Dark” covers are dominated by blacks and dark purples, blues, and grays – essentially, inky nighttime colors. Earth tone color schemes are dominated by natural blues, browns, and greens; covers denoted as photographs consist of full-color photographs without filters (including sepia and black and white filters); color schemes marked “reds & blacks” are, self-explanatorily, dominated by red and blacks. Deep solids are instances of covers dominated by a single rich, dark color (for example, hunter green, burgundy, or violet); primary colors are instances of bright, true orange, red, yellow, green, or blue, without any gradient, effect, or filter. Pastels are instances of light, pale colors, in combination or as a single color, and color schemes marked “acid greens & yellows” are dominated by bright acid green with tinges of eerie, sickly yellow. I chose these categories specifically because they

constituted the smallest number of categories that both included every cover and didn't force a cover into a category in which it didn't obviously belong. Fewer categories would have resulted in covers being squeezed into categories with which they did not share similar enough characteristics, and a greater number of categories would have resulted in such narrow definitions of categories that the data would start to lose meaning – for example, there might be only one solid dark green cover, which is meaningless on its own but which gains meaning as a data point when grouped into a category with other solid, dark-colored covers. The types of covers represented by the categories emerged organically – as I noticed more and more covers with the same distinctive acid green and yellow color scheme, for example, I developed a category to include them.

With exception of pastels and primary colors, all of these color schemes are similar in that they can all be categorized as in some way serious. A sepia color scheme gives the impression of age, time passing, and reminiscence; it is atmospheric and thoughtful. A dark color scheme gives the impression of solemnity, grimness, and sadness; earth tones suggest realism and a story grounded in a concrete universe. Reds and blacks suggest violence and brutality; acid greens and yellows are mysterious, eerie, and unsettling. Picking up one of these books, it is unlikely that a consumer would expect a lighthearted book, a silly book, or a comedy. Rather, these covers prime a potential reader for novels about serious, sad, or potentially violent topics; that is, stereotypically literary topics and topics appropriate for intellectual adult novels. These covers are not, as Figure 4 shows, dominated by bright neon colors or pinks or distinct prints like camouflage or a comic book page, all of which might suggest childishness of content or a

gimmick or, in the case of pinks and purples, a gendered “chick lit” type of novel; there is little youth or stereotypical femininity associated with the colors used.

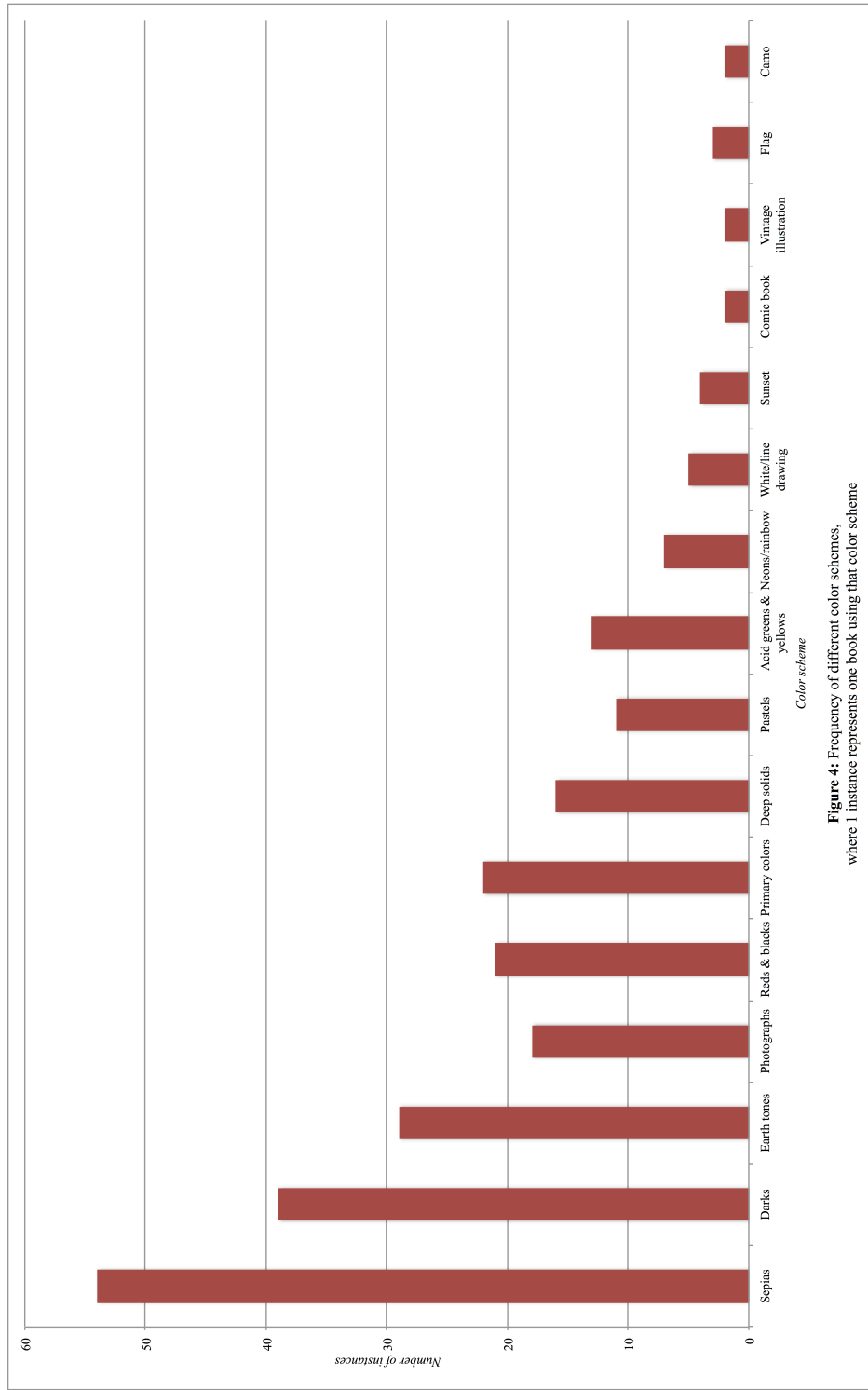


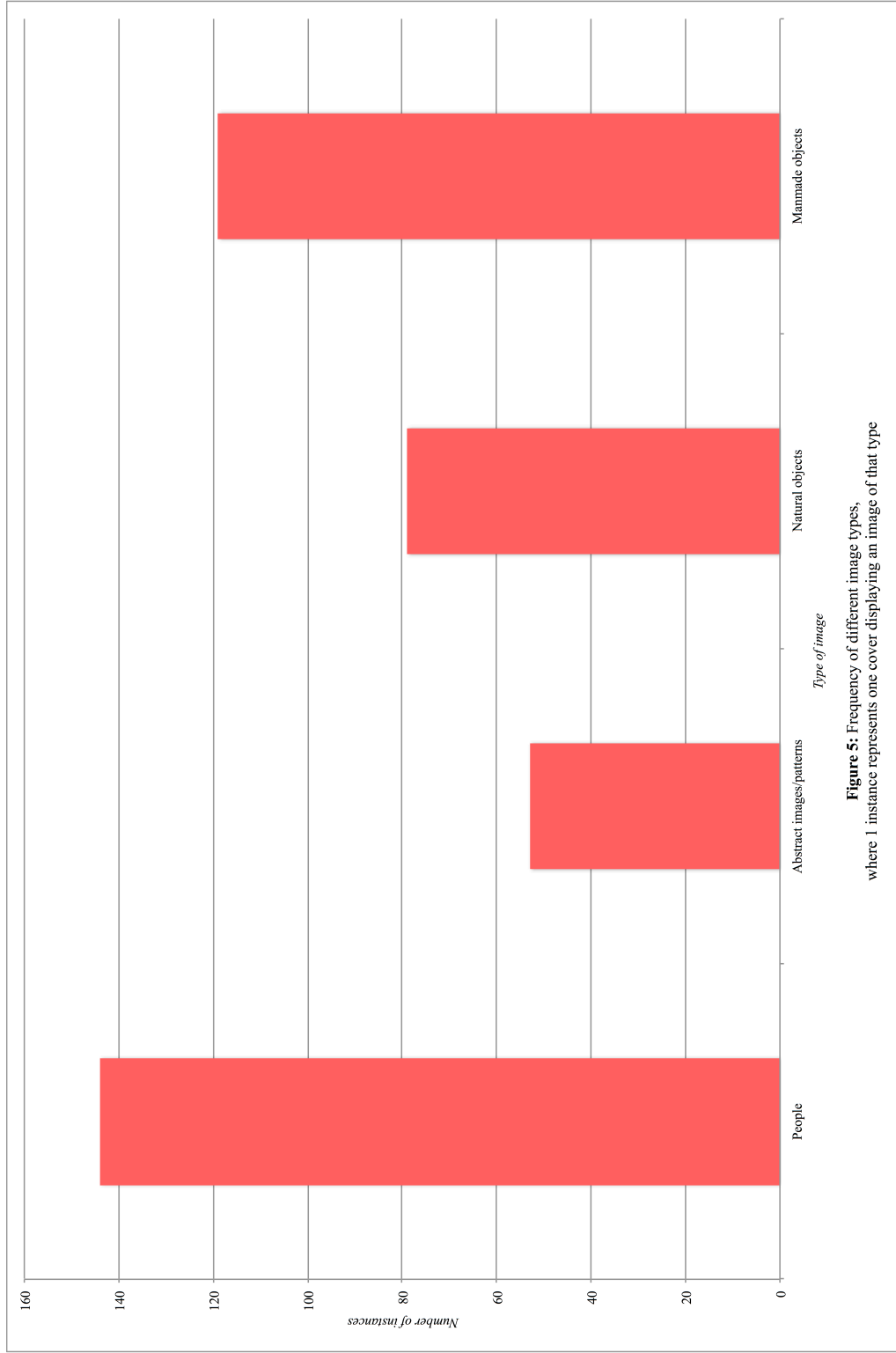
Figure 4: Frequency of different color schemes, where 1 instance represents one book using that color scheme



## Image Type

Figure 5 shows how frequently different types of images were used on the covers of the novels that won or were shortlisted; any one cover might have had more than one type of image. The most frequently recurring image type was images of people, followed closely by images of manmade or inanimate objects (objects like tables, chairs, lamps, buildings, fences, or streets); images of natural objects, like trees or oceans, were moderately common, while abstract images or patterns were the least common. These categories were developed such that the images on each cover would fit comfortably and obviously into a category, and I avoided making more finely grained categories so as to avoid creating so many categories that the data lost meaning – for example, breaking down the inanimate object group into “roads,” “tables,” etc., would have resulted in so many tiny groups of images that the data would not demonstrate trends as well.

Firstly, the prevalence of people and manmade objects suggests a focus on human interaction with the manmade world, rather than a focus on human interaction with the natural world, as well as a preoccupation with the industrial, developed world. The lack of abstract objects and patterns, on the other hand, hints at a preoccupation with realism among the novels – these covers are not depicting fantastical images or bizarre images or even nonsensical images, but almost overwhelmingly clearly identifiable images, images recognizable as realistically drawn people or buildings or roads or rivers. The consumer is primed to expect a novel about real life, about realistic people living in realistic places and having experiences that, while perhaps not common, are still possible given the constraints of our universe. Once again, the covers of the novels lead a potential reader to expect serious, realistic stories, rather than silly or fantastical ones.



**Figure 5:** Frequency of different image types, where 1 instance represents one cover displaying an image of that type

## Font Type

Figure 6 shows how often different kinds of fonts were used as the primary font on covers of novels that won or were shortlisted for the Booker Prize. For reference, a serif font is one in which the letters have small lines attached to the ends of strokes – serif fonts include fonts like Times New Roman and Garamond. Sans serif fonts are ones in which the letters do not have lines attached to the ends of strokes, such as Arial or Century Gothic. Stylized fonts are any fonts that do not fall into the serif or sans serif categories – that is, script fonts or mock-handwriting fonts, for example. Covers that use primarily serif fonts are represented by the first bar on Figure 6; covers that use primarily serif fonts are represented by the second bar; covers that use an equal combination of serif and sans serif fonts are represented by the third bar; and covers that use primarily stylized fonts are represented by the fourth bar.

Clearly, the large majority of covers used primarily serif fonts; among the covers that use serif fonts, most of those covers use all capital letters or capitals and regularly capitalized text, with few covers using only regularly capitalized text or no capitals. Stylized fonts are the second most commonly used type of font, and very few covers use only sans serif fonts or serif and sans serif fonts in combination. This further indicates the tendency of novels that win or are shortlisted for the Booker Prize to be packaged as serious books, rather than fanciful or funny, as serif fonts are often used with academic writing and print media. The insides of both fiction and non-fiction books, journals, magazines, and newspapers are typically in serif fonts, for example. Sans serif fonts, on the other hand, are often associated with digital media – text messages and websites are

often in sans serif fonts, for example. A serif font primes a consumer to expect intellectualism, authority, and permanence, more than a sans serif font might.

Stylized fonts are perhaps less suggestive of a specific trend among all Booker Prize novels, as the variety among styles and appearances means that they can be more easily customizable to suit a specific novel’s content or theme. However, script fonts may lend a certain elegance or sophistication to a cover that a serif font or a sans serif font might not.

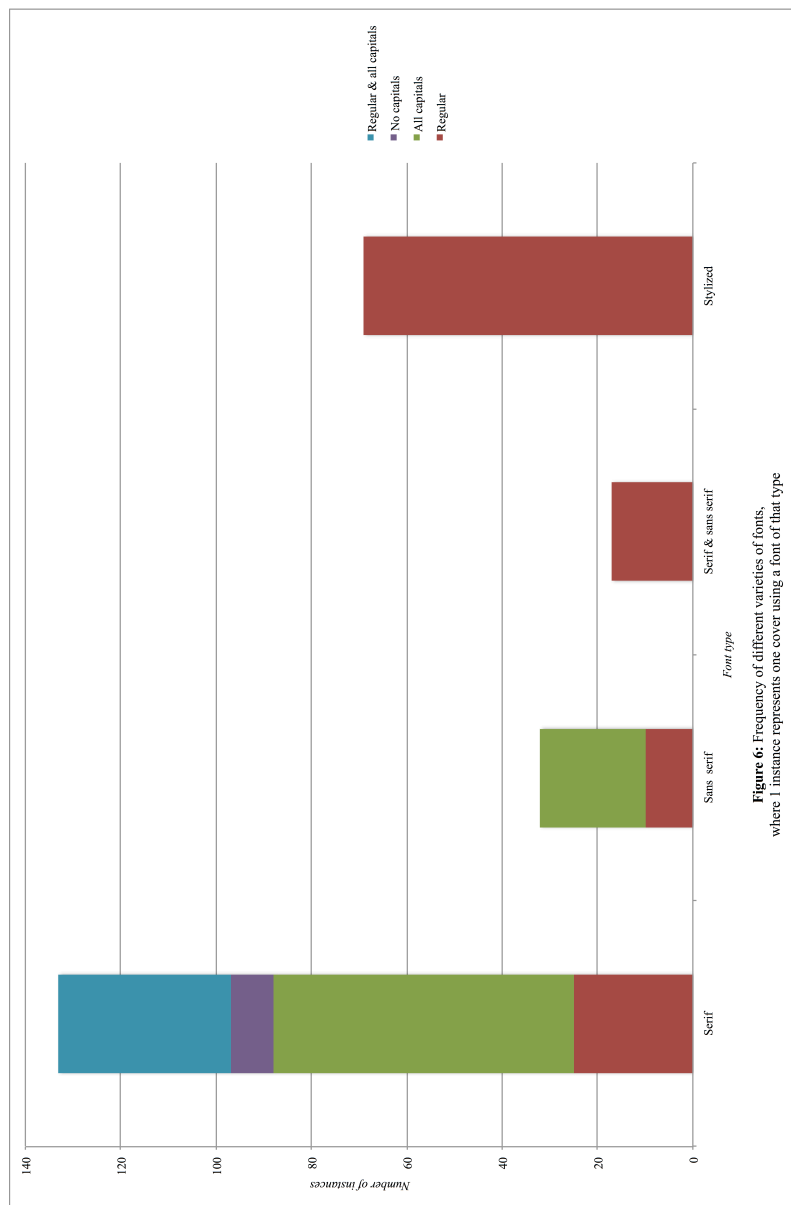


Figure 6: Frequency of different varieties of fonts, where 1 instance represents one cover using a font of that type

## *Formal Variables*

### Page Length

Figure 7 demonstrates the average page length per year of the novels that won or were shortlisted. The dotted black line is a trend line indicating that the novels have been getting longer on average over time, and the  $R^2$  value of 0.29755 indicates that approximately 30% of the variation in page length over time is explained by the linear correlational relationship between page length and time. This isn't a terribly high  $R^2$  value, but it's high enough to indicate that a trend is definitely evident and that the increase isn't purely chance or coincidence.

The colored lines in Figure 7 represent the average page lengths of novels of different genres as reported by the Book Genome Project ("The Book Genome Project"); until the mid 2000s, the average length of the shortlist was shorter than the average length of books in all genres except mystery and romance and was dramatically shorter than science fiction, fantasy, and historical fiction. Initially, the average length of the shortlist was shorter than all other genres, but this is a somewhat unreliable conclusion given that the data from the Book Genome Project includes contemporary novels as well as older novels and thus does not account for an increase in book length across genre over time. In the mid to late 2000s, the presence of several very long (400 pages or more) books on the shortlist – Hilary Mantel's *Wolf Hall* and *Bringing Up the Bodies*, as well as A. S. Byatt's *The Children's Book* and Will Self's *Umbrella* – dramatically increased the average page length per year, accounting for the sudden spike on the graph. Without this spike, the trend line might not have such a high positive slope, though there would still be a positive correlation between page length and time.

Interestingly, both *Wolf Hall* and *Bringing Up the Bodies* are historical fiction, written about Thomas Cromwell and Tudor England – most other books on the shortlist, while perhaps not set in the present, are not historical fiction in the sense that they do not tell a story about real historical characters in a specific, documented historical context, and most other books on the shortlist are nowhere near as long. Additionally, hugely long novels or series of novels are often epic fantasy or science fiction sagas – J. R. R. Tolkien’s *Lord of the Rings*, George R. R. Martin’s *A Song of Ice and Fire*, or Robert Jordan’s *A Wheel in Time*, for example – and fantasy and science fiction novels are also decidedly genre novels. Very few novels on the shortlist qualify as fantasy or science fiction, either, in that they are set on Earth as we know it and do not contain fantastical, supernatural, or magical elements. This relationship between length and genre suggests a predisposition of the novels that win or are shortlisted for the Booker Prize to avoid length at the same time as avoiding an easily niched genre classification – in its shortness, a Booker Prize novel further establishes itself as serious and intellectual by avoiding association with the length and density of a fantasy, science fiction, or historical fiction novel, which, as a genre novel, might be associated with a lack of intellectualism and importance.

Moving into analyzing individual texts, we can see an interesting relationship between page length and the type of story told. Both *On Chesil Beach* and *The Remains of the Day* are fairly slim, at 203 pages and 245 pages respectively; *Cat’s Eye*, on the other hand, is 462 pages long, and is a much different type of story in many respects. *On Chesil Beach* tells the story of Edward and Florence, a newly married couple in 1962, who go on honeymoon to Chesil Beach and, due to sexual repression and inexperience,

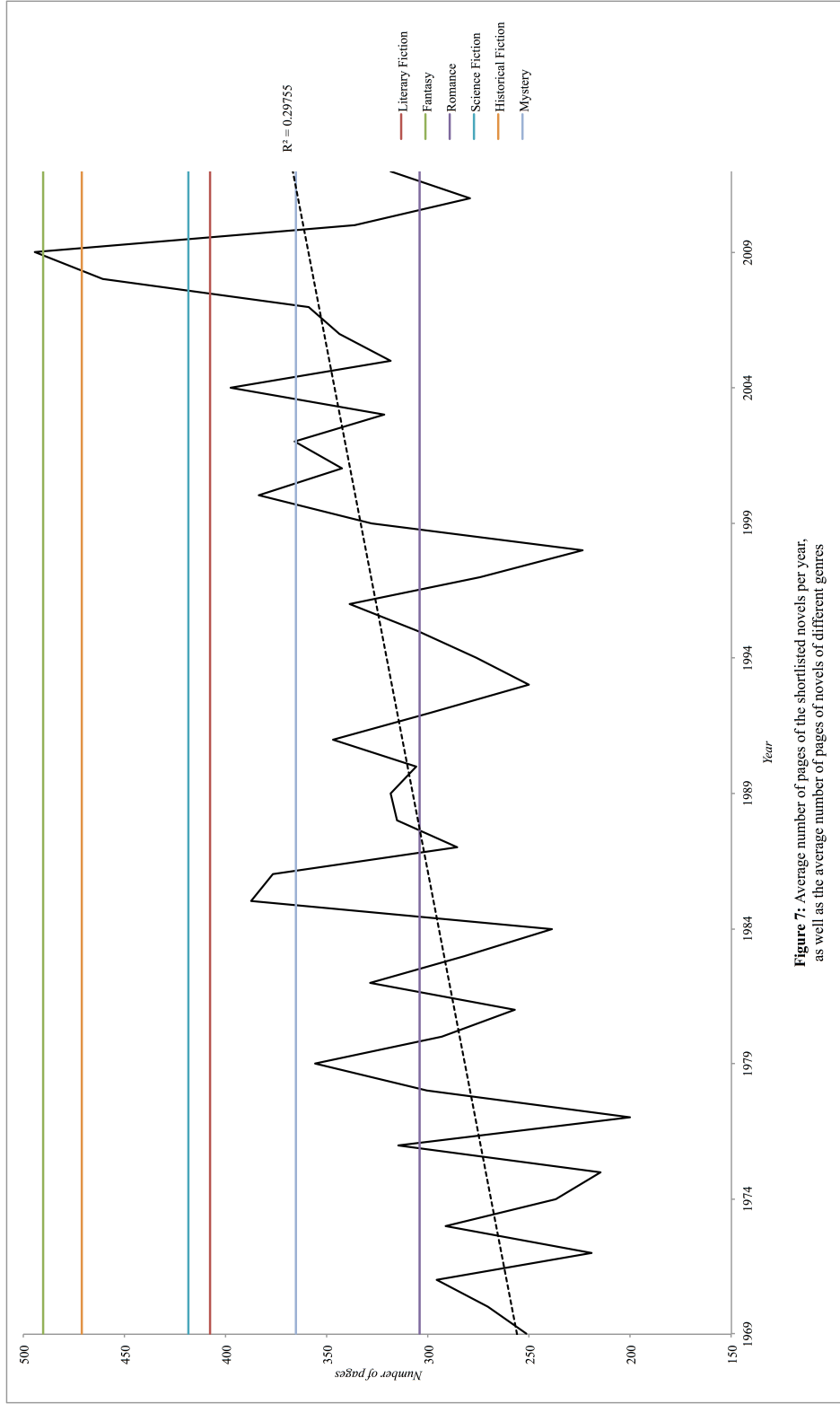
end up destroying their newborn marriage; the story focuses largely on the couple's experience on Chesil Beach, with flashbacks and flashforwards to inform the action occurring in the present. *The Remains of the Day* is the story of Stevens, an aging English butler of an old English country estate, who goes on a road trip at the suggestion of his new American employer; Stevens reflects on his previous employer, the ruined and now-dead Lord Darlington, to whom he was fiercely loyal, and eventually visits his old colleague, Miss Kenton, for whom he once had feelings. Again, while the narrative uses flashback to inform the present-time narrative, the majority of the action occurs in that present-time; as with *On Chesil Beach*, the novel's climax and primary tensions are being dealt with in the present and are informed by the tensions and actions of the past. Both these novels, as a result, are brief and almost impressionistic; the present-time action takes place in the span of Edward and Florence's brief one-night honeymoon or Stevens' brief several-day drive across England. The novels are short because they take place in a matter of days; while the backstories of the characters take place across years, these backstories enter the narrative as short, vivid scenes or obscured, brief flashbacks narrated by the characters.

*Cat's Eye*, on the other hand, rather than grounding itself primarily in a brief span of present-time, spans the entire lifetime of its narrator, Elaine, a relatively successful painter who returns to her hometown of Toronto for a retrospective of her work and eventually confronts the ghosts and traumas of her childhood. Each era of Elaine's past, from early girlhood to puberty to young adulthood, is given equal weight in terms of pages, with the first 250 pages of the novel dealing primarily with Elaine's life prior to age thirteen. Elaine's present, rather than serving as the locus of action that is informed

by the past, serves instead as a frame story that gives context to the dramas and characters of the past. The novel reads much less like an impressionistic glimpse into a few days of a character's life and much more like a detailed saga of a character's entire existence and, as such, it is more than twice as long as *On Chesil Beach* and *The Remains of the Day*. *Cat's Eye* is also written by a writer famous for her science fiction and dystopian novels, such as *The Handmaid's Tale* and *Oryx and Crake*, and includes a few supernatural or dreamlike elements itself, like Elaine's vision of a Virgin Mary holding a blazing cat's eye marble who comes to rescue her from an icy ravine after her friends abandon her. This suggests that perhaps some of the seemingly anomalous elements of *Cat's Eye* are related; its saga-like nature results in a page count more characteristic of typically saga-driven genres like fantasy and science fiction, while the crisp impressionistic qualities of *Chesil Beach* and *Remains* result in page counts much more typical of a Booker Prize nominee.

Ultimately, looking at these three novels, it is clear that page count is not simply an unimportant measure of how long a book is; it reveals something important about what kind of story a novel may be telling. Long novels have the space to be sagas and epics, to be stories that span years or decades and give attention to each of those years; short novels have the space for stories detailing the intricacies of a few days or a series of brief, vivid flashbacks. The small page count of the average Booker Prize nominee indicates that the nominees are more likely to tell the later kind of story, with long, saga-like novels being atypical and rare.





**Figure 7:** Average number of pages of the shortlisted novels per year, as well as the average number of pages of novels of different genres

## *Content-based Metadata*

### Geographic Setting

Figure 8, similar to Figure 3, shows where novels that won or were shortlisted for the Booker Prize were set; if an area on the map is shaded in, at least one book included that location as one of its primary settings, and the darker the shading, the more books that included that location as a primary setting. For the purposes of this project, a location is a primary setting if a large portion of a book takes place there; a place that is visited briefly or a place that is merely mentioned does not count, and if a book contains an at-sea or off-world setting, that setting isn't included in this particular figure.

Apart from the United Kingdom and the Republic of Ireland, which feature as settings in so many novels that they require a separate figure (45% of all settings), the most common settings are France (6% of total settings), India (6%), the continent of Africa as a whole (7%), Australia and New Zealand (3%), Canada (3%), the United States (7%), Italy (3%), and Germany (3%). In total, this accounts for approximately 83% of all settings, leaving only 17% left to the entire rest of the globe, including all of Asia except India and all of South America. All of these settings are strongly tied to Britain and England – Australia, Canada, and New Zealand are all part of the Commonwealth, for example. India and much of Africa have strong ties to British Imperialism and colonization. France, Italy, and Germany are all Western European countries with strong wartime ties to England, especially as all four countries were intimately entwined during World War II. The United States, as a former colony of the British Empire and a close political ally currently, is also strongly connected tied to the UK and England.

Clearly, even if the novels are not set in the United Kingdom or the Republic of Ireland themselves, they are primarily set in locations strongly tied to Englishness or Britishness, especially English or British nationalism in colonial or wartime contexts. These 254 novels are not telling stories about countries all over the world – they are telling stories about the world as it relates to England and the United Kingdom and creating a universe with England at the center and all other countries functioning merely as actors in England’s story. Even if they aren’t telling stories that glorify or promote Englishness or English colonialism, they will still be representing the world as it relates to England, even antagonistically. The world as seen in the corpus of Booker Prize winners and shortlisters is an England-centered, England-focused world.

The inset to Figure 8 provides further detail about this English centrality. London is, by far, the most popular English setting – it accounts for approximately 40% of all UK settings and 18% of total settings. Middle England comes in second, at approximately 17% of all UK settings, followed by Northern Ireland and the Republic of Ireland at 13%, Southern and Northern England at approximately 9% each, Scotland at approximately 5%, and Wales, the Orkney Islands, and the Isle of Wight at under 5% each. No books were set in Cornwall at all. Once again, London, the epicenter of Englishness and English national identity, accounts for the plurality of English settings, further confirming the emphasis on English nationalism and identity among the settings of the Booker Prize novels; the rest of England is also well-represented, as is Ireland, which has historically been in conflict with England and English colonialism, but places in the United Kingdom less associated with English identity, like Scotland and Cornwall, are not well-represented.

Clearly, setting in these novels is more than a random choice and more than a simple location for action to take place – as is evident from examining individual texts, setting functions not only to direct the novels’ perspectives but also to determine the character of the novels. *On Chesil Beach* and *The Remains of the Day* are both intimately tied to the English landscape; in *On Chesil Beach*, for instance, the eponymous beach is of “infinite shingle” (McEwan 4), isolated and bordering gray, lapping water. The weather is “not raining, but nor [is] it quite warm enough” (McEwan 4). This serves as an effective backdrop for the repressive interiority and stifling politeness of its protagonists, who are separated from one another by an ocean of silence and who are as English as a beach that is not sunny but not quite raining either. The English seaside, iconic as it is, functions to prime the reader to expect Englishness, to expect politeness and quietness and isolation. Similarly, in *The Remains of the Day*, Stevens drives across England, from his country estate to Salisbury to small towns and Dorset and Somerset and Devon and finally to a small town in Cornwall; in each location, he encounters the kind of rural, pastoral landscape that represents the fading old England he knew so well, the fading England that is giving away to urban modernity. The country estate, the small village – these are remnants of a country that no longer exists. Again, landscape and setting function to convey a kind of national and temporal identity, and if one setting is overrepresented – in this case, England – this national identity is overrepresented as well.

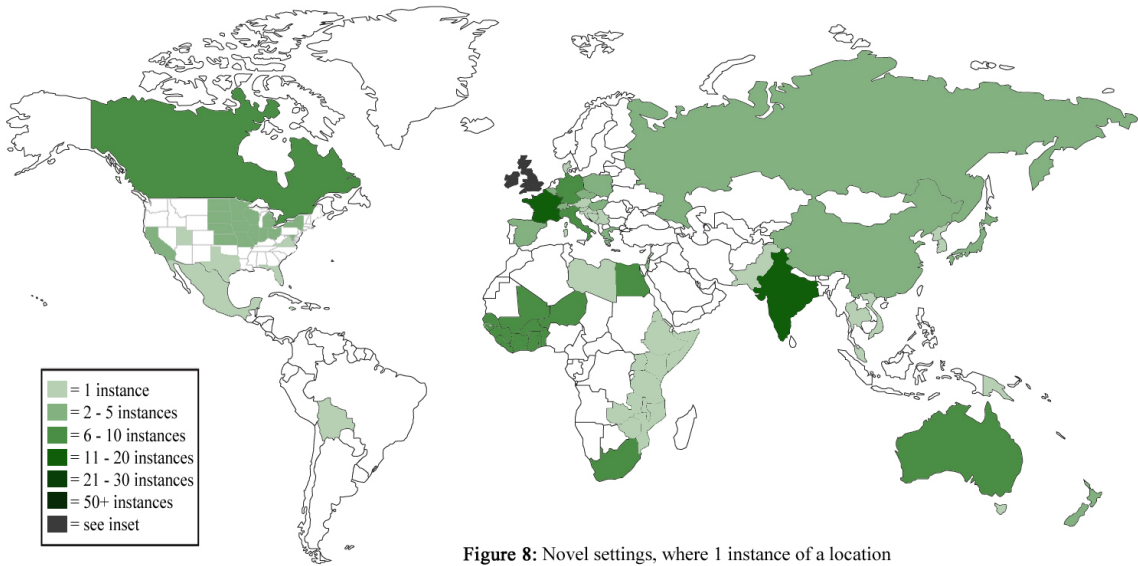


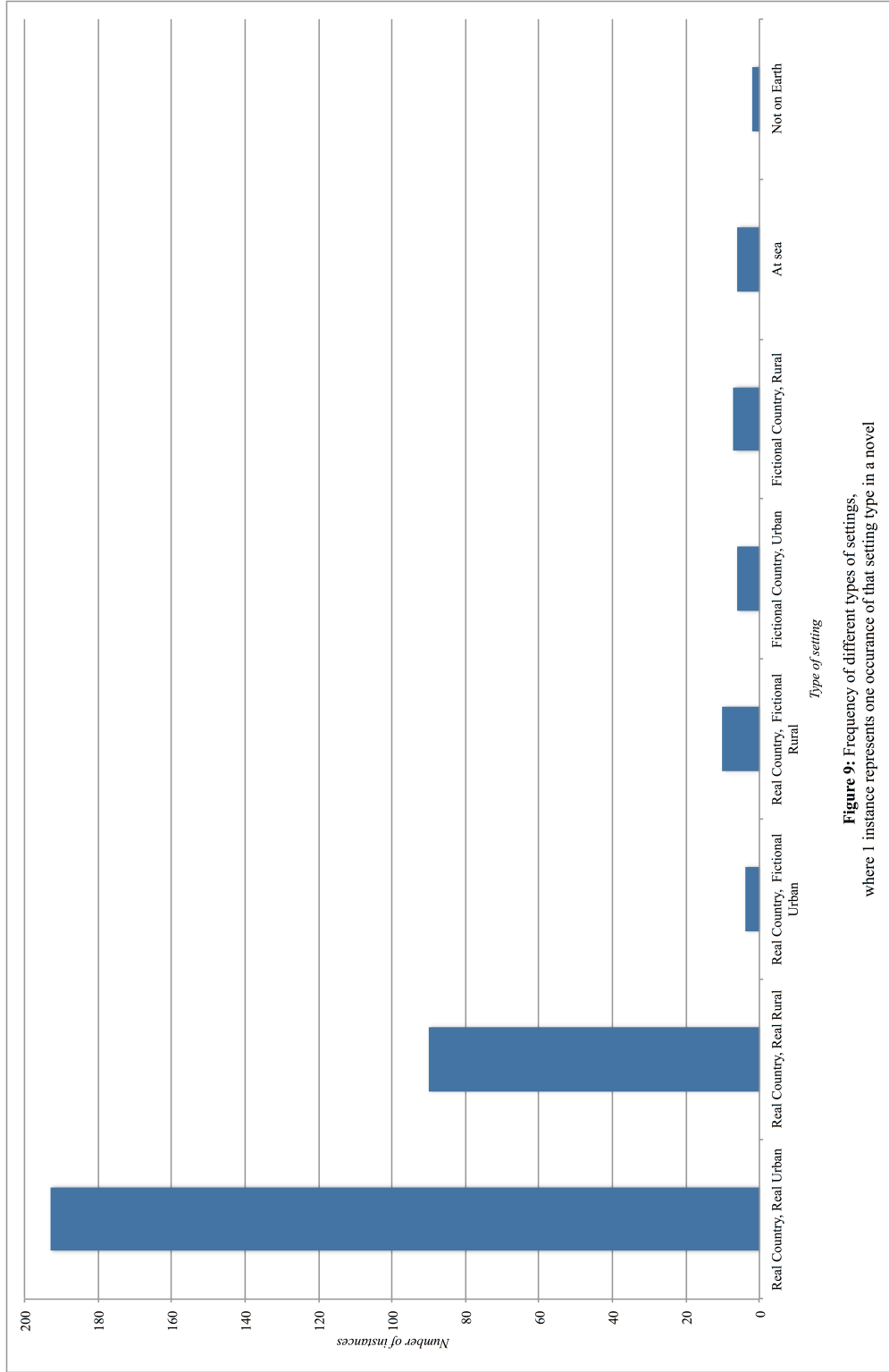
Figure 8: Novel settings, where 1 instance of a location represents one appearance of that location as a setting in a novel



## Setting Type

Figure 9 demonstrates how frequently different types of settings, rather than specific geographical locations, are used as primary settings among the corpus of Booker Prize novels. Real cities in real countries make up the majority of settings, followed by real rural areas in real countries; very few novels use explicitly fictional cities or explicitly fictional rural areas, even in real countries, and even fewer use fictional cities or rural areas in fictional countries. Hardly any novels are set at sea or on a world other than Earth, but a few novels do include one or the other of these two categories of settings.

The fact that the vast majority of the corpus is set in real cities or real rural locations goes to reinforce the idea that the Booker Prize novels tend to focus on explicit reality, rather than the imagined or supernatural or otherworldly – they are concerned with stories that take place in a world recognizably and identifiably ours, rather than stories that take place in imagined places. The fact that so few novels take place off-world also goes to reinforce the idea of genre novel avoidance – Booker Prize novels are fiction, but they aren't science fiction or fantasy and they are explicitly and intricately tied to the real world in a way that a science fiction or fantasy novel might not be. This suggests that value is placed on realism across the corpus. Books that are literarily valuable are books that focus on, function within, and offer insight on the real world; they deal with grim realities, rather than pretend places, and this resistance towards the completely imagined is part of what qualifies them to be nominated for the Booker Prize.



**Figure 9:** Frequency of different types of settings, where 1 instance represents one occurrence of that setting type in a novel

## Time Period

Figure 10 shows how frequently certain time periods appear as primary time periods among novels that won or were shortlisted for the Booker Prize. Once again, a primary time period is one in which a large portion of a novel is set, not one that only appears briefly or is merely mentioned. Additionally, Figure 9 represents the totality of special time periods – unless a novel’s time period is represented here, the novel is set entirely post-1950.

Of the 318 time periods included in the Booker Prize corpus, only 90 time periods (28%) occur pre-1950; this suggests a focus on the present day, modernity, and the ways in which modernity interacts with the past. This complements the novels’ focus on the real world quite well – these books are focused on world we know and the time period we know, and imagining an alternate world or a time period other than the current one is a much less common practice among nominated novels than operating within a known and verifiable framework. Additionally, this lack of historical novels suggests potential avoidance of historical fiction, as novels set in the past are, obviously, far more likely to involve specific historical events or historical figures than are novels set post-1950.

Of those 90 pre-1950 time periods, the most frequently occurring is the 1837 to 1901 time period, followed by 1939 to 1945, followed by 1945 to 1950, 1914 to 1919, 1901 to 1913, and 1930 to 1938. All of these time periods are closely connected to a historical era involving significant upheaval or national identity development for England and/or the United Kingdom. 1837 to 1901 represents England’s Victorian Era, which was a post-Industrial Revolution period of English confidence and the refinement and proliferation of English culture and sensibilities. It encompasses the peak of the British



Empire; the writing careers of quintessentially English authors like Charles Dickens, the Brontës, Thomas Hardy, and George Eliot; the codification of the sexual and social restraint we contemporarily associate with the period; massive population and architectural growth; and a long period of military peace (the Pax Britannica). A focus on the Victorian Era indicates a focus on a time period linked to the flourishing and development of a successful, superior English identity.

All five remaining common time periods are wartime periods or periods immediately preceding or following a major war – specifically, World War I and World War II, with time periods related to World War II being more common than time periods relating to World War I. Both World War I and World War II were time periods of social and political upheaval and change and, since both wars involved nation versus nation clashes, the formation and solidification of English national identity as contrasted with external national identities. In World War II especially, England as a nation appealed to its citizens' senses of patriotism and national pride to support the war effort, and the disproportionate number of novels set during these time periods indicates a focus on this kind of national identity and national crisis – again, a focus on Englishness. In contrast to the military peace present in the Victorian Era, however, the violence and brutality fundamental in war also suggests a preoccupation with pain, violence, tragedy, and cruelty, as well as the high stakes of an individual facing a high probability of death.

Applying Figure 10 to actual texts, both *The Remains of the Day* and *Cat's Eye* incorporate some aspect of the causes or aftermath of World War II, but *The Remains of the Day* offers a much clearer picture of how WWII can function as a temporal setting to create a specific atmosphere and tone and to establish a specific transitional space

between a kind of antiquity and a kind of modernity. In *The Remains of the Day*, which is set in 1956 but which incorporates a good deal of reminiscing and flashing back to the 1930s and 1940s, it gradually becomes clear that Lord Darlington was an extreme right-wing conservative during WWII, with ties to German Nazi organizations, and that part of Lord Darlington's ruin was the result of his political ties. Stevens even reveals that he himself once fired two Jewish servants on Lord Darlington's orders. Stevens, now employed by an American in the wake of Lord Darlington's death, is left in the liminal space between England as it existed prior to WWII and England as it began to exist afterward; WWII acts as the dividing line between two very different Englands, and Stevens is stuck in the old one as the rest of the country re-crafts itself in the wake of the war. As Lord Darlington puts it, "We [England] are always the last, Stevens. Always the last to be clinging on to outmoded systems. But sooner or later, we'll need to face up to the facts. Democracy is something for a bygone era. The world's far too complicated a place now" (Ishiguro 198); of course, Lord Darlington used this sentiment as a justification for siding with the Nazis, but the sentiment itself is not entirely wrong. In this novel, WWII represents a split between an old era and a new one, both politically and socially – England as it existed previously, slow to change and clinging tightly to old traditions, could not exist any longer as the world was launched into a kind of chaos for which it could never have prepared itself, and anyone still subscribing to an older version of Englishness found themselves suddenly outmoded, dated, and out of touch.

Ultimately, both Figure 10 and *The Remains of the Day* demonstrate that the corpus of Booker Prize novels focus either the present and ways in which English identity has been restructured and called into question since 1950 or a time in the past when

Englishness was challenged, broken, or reaffirmed. They focus on solemn time periods, time periods of violence or pain or tumult, and they deal with how individuals interact with the darkness surrounding them. A novel's temporal setting, in this case, is important because the setting selected guides and shapes the way nationality and, in particular, Englishness, function in the text, as well as how the past relates to the present.

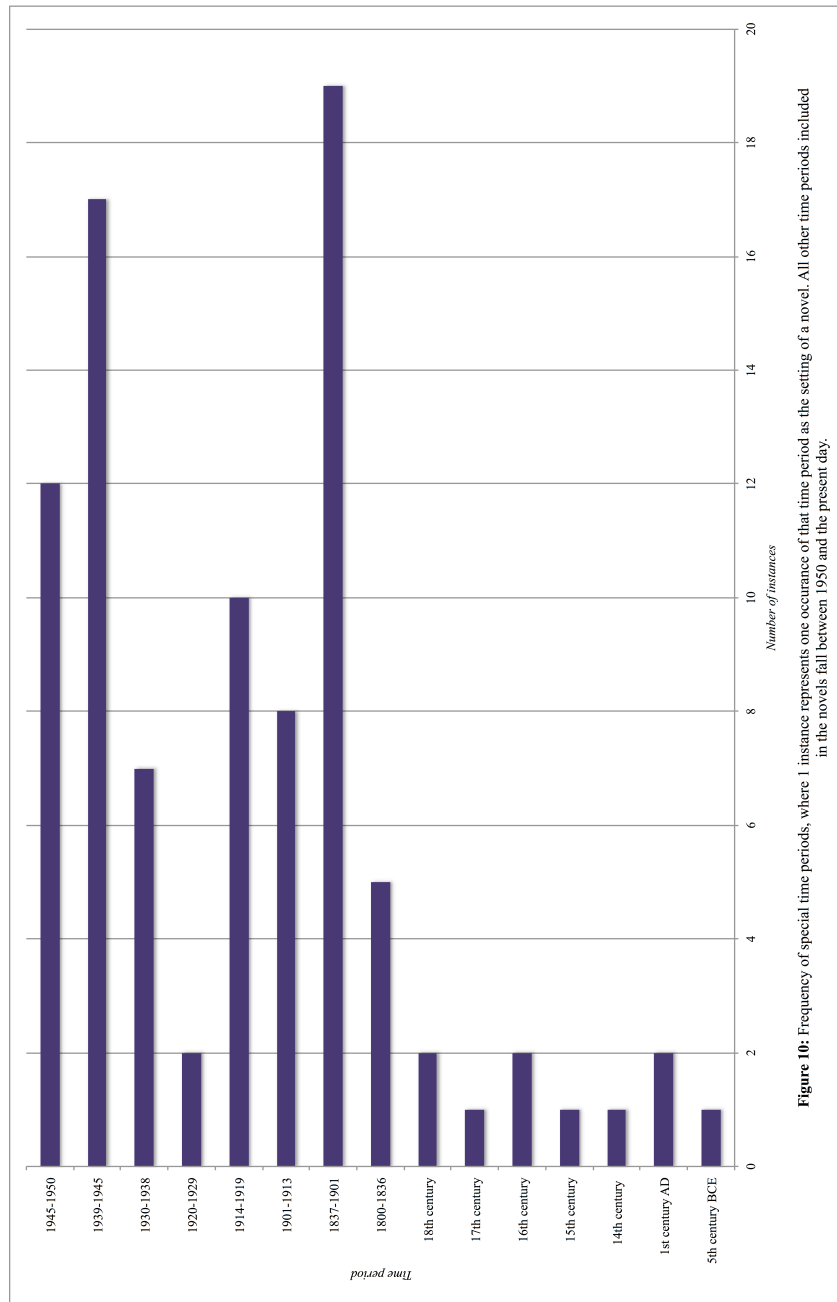


Figure 10: Frequency of special time periods, where 1 instance represents one occurrence of that time period as the setting of a novel. All other time periods included in the novels fall between 1950 and the present day.

## Protagonist Type

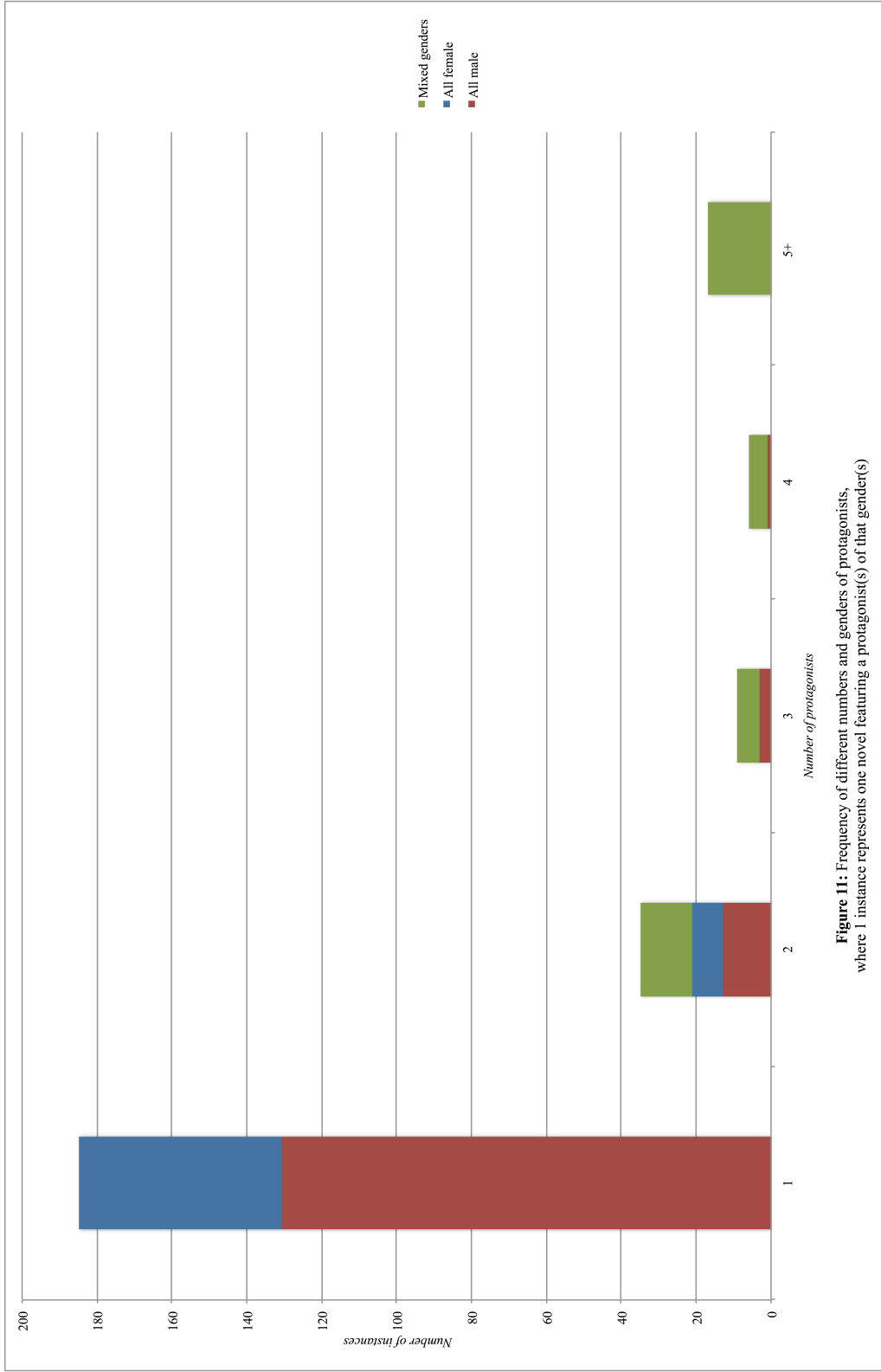
Both Figure 11 and Figure 12 break down and represent the kinds of protagonists featured in the corpus of Booker Prize novels; in this context, a protagonist is a point-of-view character or a character whose experiences and thoughts are central to the narration or plot. A character who simply appears in the book frequently does not count, because this character's perceptions of and reactions to the world may not be the perceptions and experiences driving the novel.

Figure 11 depicts the number of novels featuring different numbers of protagonists of different genders – that is, novels with only one male protagonist, novels with two female protagonists, novels with one male and one female protagonist, etc. The vast majority of novels feature only one protagonist, and the majority of novels with only one protagonist feature a single male protagonist rather than a single female protagonist. Very few novels feature two protagonists, though more novels feature two protagonists than three protagonists, four protagonists, or ensemble casts (five or more protagonists); of the novels that feature two protagonists, these novels are equally split among two men, two women, and one man and one woman, and of the novels featuring three or more protagonists, most feature combinations of both male and female protagonists.

The extreme bias towards singular male protagonists among the corpus of Booker Prize winners and shortlisters suggests a bias towards documenting the world from a single, masculine viewpoint – these novels are not about the experiences of communities or groups of people so much as they are about the interactions of singular individuals with the world around them. A single protagonist indicates a focus on individuality, internal experience, and isolation, as a story with one protagonist is much more likely to

depict the internal monologue of one person against the world, to the exclusion of all other voices, than is a story with multiple protagonists; additionally, communal narrative may connote as more female, while individuated narrative may connote as more male, which may be another level of subtler masculinity within the corpus.

Additionally, the fact that there are more male protagonists than there are male novelists means that some female novelists are writing about men – specifically, of the 131 novels with single male protagonists, approximately 22% of those novels were written by women, and of the 53 novels with single female protagonists, approximately 20% were written by men. What this indicates is that because there are more male writers than female writers, because men are much more likely to write about men than about women, and because some female writers are writing about men rather than women, the perspective offered by the corpus of novels is overwhelmingly skewed towards the masculine.



**Figure 11:** Frequency of different numbers and genders of protagonists, where 1 instance represents one novel featuring a protagonist(s) of that gender(s)

## Male Protagonists

Figure 12 shows the percentage of shortlisted or winning novels per year that featured at least one male protagonist. The red line marks fifty percent (half the books in that year had at least one male protagonist), and the blue line is a trend line.

Firstly, it is clear that the majority of novels every year have at least one male protagonist. There are only four instances when the percent of books with at least one male protagonist dips below fifty percent, and in the majority of years, the percentage is seventy or above. Additionally, while there are multiple instances of all the novels in a given year featuring at least one male protagonist, there are zero instances of none of the novels in a given year featuring at least one male protagonist – that is, from 1969 to 2012, there has never been a shortlist of entirely female characters, but there are have been multiple shortlists of entirely male characters or male and female characters with female characters in co-protagonist roles only.

The blue trend line slopes upward, indicating that the percentage of novels on the shortlist featuring at least one male protagonist is increasing; this is opposite the percentage of female authors on the shortlist per year, which is decreasing. The  $R^2$  value here is fairly low – only 5% of the variation is accounted for by the linear relationship between male protagonists and time – but an upward trend is still discernible.

This masculine focus is also evident in *On Chesil Beach*, *The Remains of the Day*, and *Cat's Eye*, though *Cat's Eye*, the only one of the three written by a woman, is the least biased towards masculinity and masculine perspectives. *On Chesil Beach* features one male protagonist and one female protagonist and, as a result, Florence, the female protagonist, is defined almost entirely by her relationship to her fiancé Edward, to men in

general, and to sex. Florence's primary conflict is the visceral disgust she feels in the face of sex, a disgust she explains as "a dry physical sensation of tight shrinking, general revulsion at what she might be asked to do, shame at the prospect of disappointing him, and of being revealed as a fraud" (McEwan 103); indeed, the collapse of their marriage hinges on the moment Edward orgasms unexpectedly on a wildly uncomfortable Florence, who promptly runs from the room and onto the eponymous beach. Both characters are "too polite, too constrained, too timorous, they went around each on tiptoes, murmuring, whispering, defending, agreeing" (McEwan 180), and they are unable to communicate or salvage their relationship. The novel itself treats Florence's disgust and anguish with just as much sympathy and compassion as Edward's humiliation and confusion, and attributes their lack of communication to 1960s English identity rather than to a failing on either character's part, but it is still true that Florence does not have much of a self outside of her relationship to sex, marriage, and the male body.

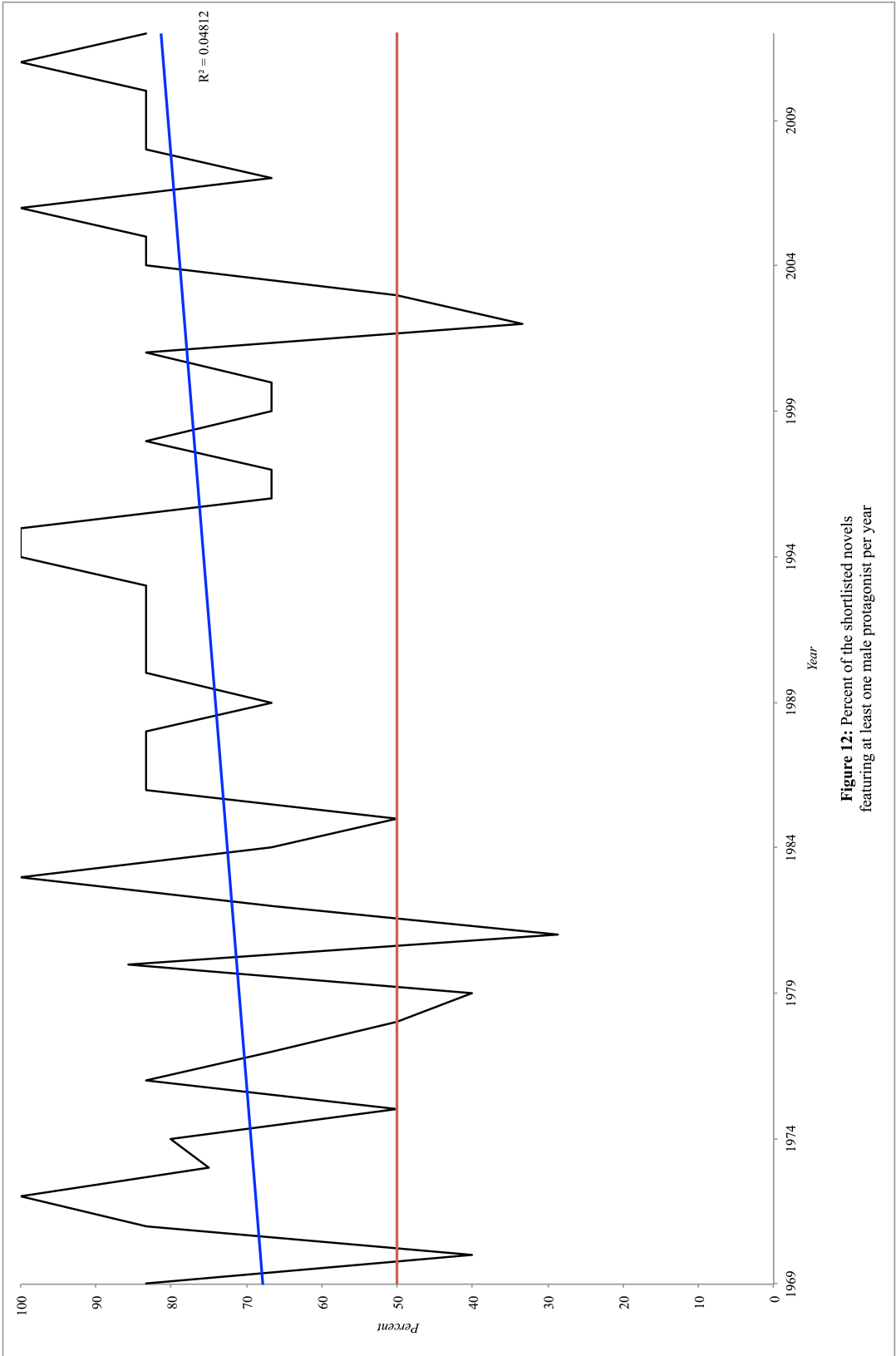
Unlike *On Chesil Beach*, *The Remains of the Day* does not have a female narrator at all; the entire novel is told from Stevens's perspective, in a very close first person, and the only female character who plays any important role is Miss Kenton, the housekeeper who was willing to question Stevens's devotion to Lord Darlington and with whom Stevens was somewhat in love. To Miss Kenton's credit, she was willing to disagree with Stevens when he was wrong, but she exists in the story only as filtered through Stevens's extremely unreliable, outdated, masculine perspective. She functions as a potential love interest Stevens is too repressed to pursue and then, at the end of the novel, as Mrs Benn, a woman who married someone else and a woman Stevens could never have. She is a foil for Stevens's stifling Englishness and overwhelming sense of professional duty, throwing



Stevens's character into relief even as she herself fails to exist in her own right. Stevens's primary devotion and affection is always directed towards Lord Darlington, even though Lord Darlington does not exist in the present time of the narrative; even when the second male character is dead, the most complex relationship is still between two men. No two women have relationships with each other, and the women characters who are present function as actors in relationships with men.

Margaret Atwood's *Cat's Eye* is the only one of the three individual texts that was written by a woman, and it is also the only one of the three featuring a single female protagonist. The entire story is told in first person from Elaine's point of view, and the relationship that pervades the text is Elaine's relationship with her childhood friend Cordelia, a relationship that continues into puberty and young adulthood. In part because Elaine was tormented and bullied by her female friends as a child, Cordelia is the character that haunts Elaine through middle age and that Elaine imagines seeing all over Toronto when she returns. Elaine does not regret her failed relationship with her husband nearly so much as the time she never had with Cordelia, who was committed to a mental institution as a young woman and who Elaine never saw again; as Elaine puts it in the final lines of the text, "This is what I miss, Cordelia: not something that's gone, but something that will never happen. Two old women giggling over their tea" (Atwood 462). Although Elaine's character is influenced largely by how she interacts with men – she grows up confused by femininity and much more at ease in the company of boys, and spends her adulthood in and out of sexual relationships with men while internally deriding her few female friends – the relationships that dominate the novel are relationships between girls and women, even if these relationships are often antagonistic.

Because the only text that features primarily complicated, intricate relationships between women is also the only text written by a woman, this suggests that the voices of female authors may indeed change the perspectives and character of the corpus. More female authors may result in more complicated female characters and female characters who exist outside of relationships with men. Additionally, between the proliferation and increase of both male protagonists and male writers, it is clear that the Booker Prize shortlisters are once again strongly dominated by male characters and male voices – when female voices are included, they are typically included alongside male voices, as a complement rather than a replacement. Rarely are the majority of novels written by women or featuring mostly women – stories about communal female experience are barely represented at all, in favor of stories about singular, individual male experience and males interacting with females. Female characters are part of male characters’ stories, or their stories revolve around men – rarely do female characters participate fully in each other’s stories and comprise a full community of characters all on their own.



**Figure 12:** Percent of the shortlisted novels featuring at least one male protagonist per year

## *Structural Variables*

### Keyword Analysis

Figure 13 demonstrates how many words each category of adjectives ended up containing, using the categories created after extracting the adjectives from the blurbs accompanying the novels in the Booker Prize online archive.

The most common type of adjective is meta adjectives – these are adjectives used in the blurb to describe the writing style, the author, or the book as a whole, rather than aspects of the book’s contents, plot, or characters. For example, in the phrase “with haunting prose, Author X writes...” haunting is a meta adjective, but in the phrase “Character X suddenly finds himself privy to a haunting family secret,” it would not be. Meta adjectives are categorized differently because while they may describe the themes or content of the book, they may also describe things irrelevant to the themes or content of the book, like the quality of the author’s vision or prose – only adjectives that described the plot, characters, or other content of the novels at hand were included in the non-meta adjective categories.

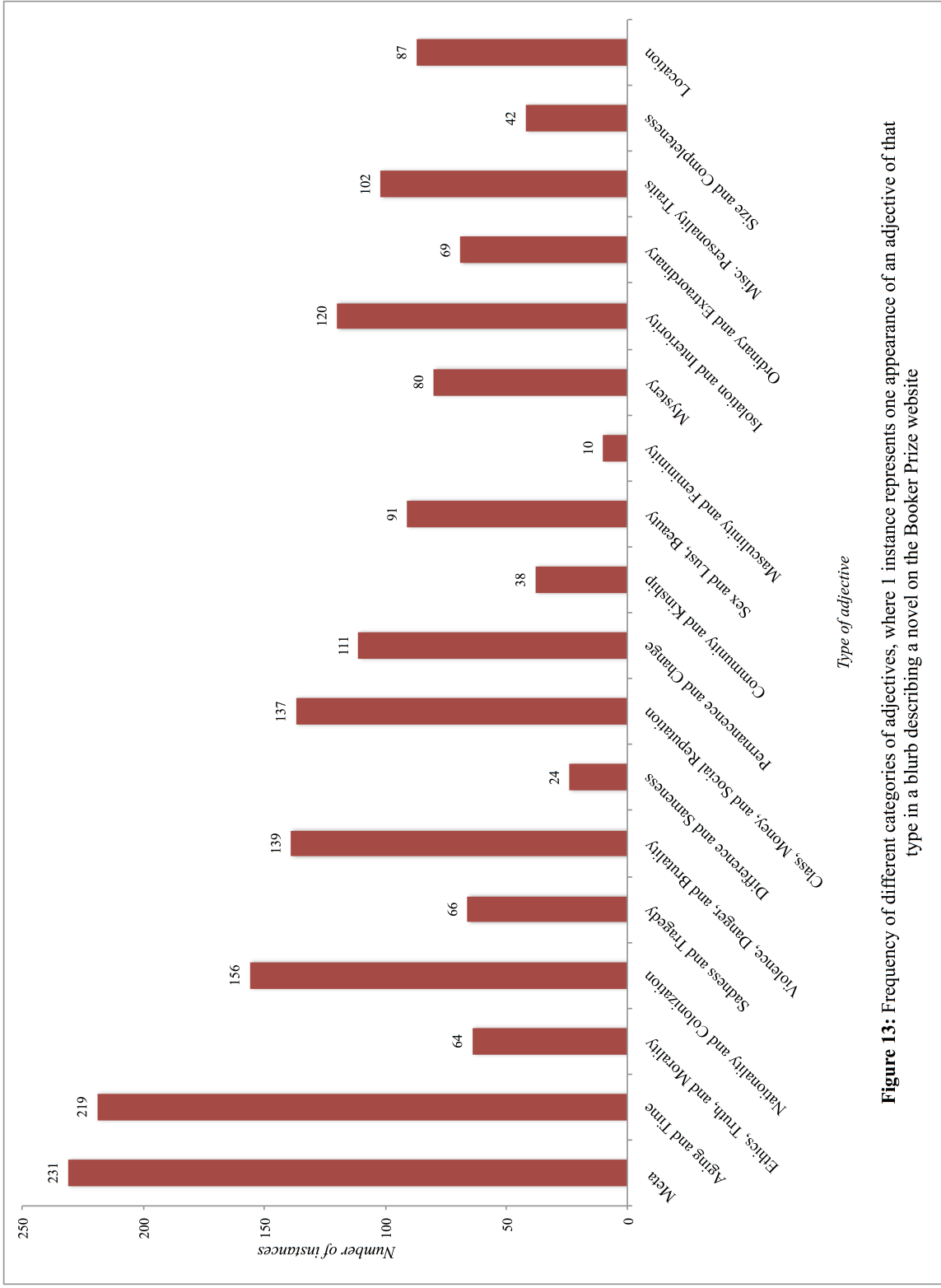
After meta adjectives, by far the most common type of adjective is adjectives concerning aging and time – any adjectives describing youth, old age, decay, nostalgia, or time period, for example, fell into this category. This suggests a thematic preoccupation among the novels with youthfulness, the transition from youth into age, and the observance of the effects of time on the world around the characters. The next three most common types of adjectives, all of which are approximately equivalent in commonality, are adjectives concerning nationality and colonization, adjectives concerning violence, danger, and brutality, and adjectives concerning class, money, and social reputation.

These three categories go to support some of what I have already found – the Booker Prize novels are focused on nationality and national identity, as well as colonial identity, as I observed by examining where the novels were set and in what time periods they were set. The novels are also concerned with violence and darkness and the nasty, ugly elements of human nature, which complements the solemnity and seriousness suggested by their covers and the proliferation of wartime settings. Finally, a focus on class and social reputation indicates that these novels focus largely on how characters interact with the material, human-made world around them, perhaps more than the natural world.

The next most common categories are adjectives concerning permanence and change and adjectives concerning isolation and interiority. A focus on permanence and change – a category including adjectives like “sudden” or “inevitable” or “unchanging” – suggests, unsurprisingly, that these are novels that often rely on surprise, action, and change to sustain their plots, as do many stories. More interestingly, however, a focus on isolation and interiority suggests that these novels are dealing with individual characters struggling within themselves and interacting with themselves, disconnected from other characters and from the world around them, facing life alone. This complements my observation that the novels feature largely single protagonists – again, these are not novels about large casts of characters or communities but novels about individuals and individual crises.

Moderately common adjective categories include adjectives concerning sex, lust, and physical beauty, adjectives concerning mystery and the unknown, adjectives concerning location, and adjectives concerning miscellaneous personality traits that described character but didn't fall into one of the other categories. The least common

adjectives concerned ethics, truth, and morality; difference and sameness; community and kinship; masculinity and femininity; and size and completeness. The rarity of some of these adjectives is not particularly telling – using a pronoun can depict masculinity just as easily as using the word “male,” for example, and the idea that two things are different or the same can be conveyed without the words “different” or “similar” – but the rarity of adjectives concerning morality and community is, if not surprising, noteworthy. This suggests that these novels are not preoccupied with characters making traditionally moral choices or engaging in ethical debates with themselves – rather, these novels seem to focus on the immoral or amoral brutality and cruelty of the world. Additionally, these novels continue to focus on the individual, rather than the community – family, partnership, and togetherness are less important than loneliness, separation, and personal struggles.



**Figure 13:** Frequency of different categories of adjectives, where 1 instance represents one appearance of an adjective of that type in a blurb describing a novel on the Booker Prize website

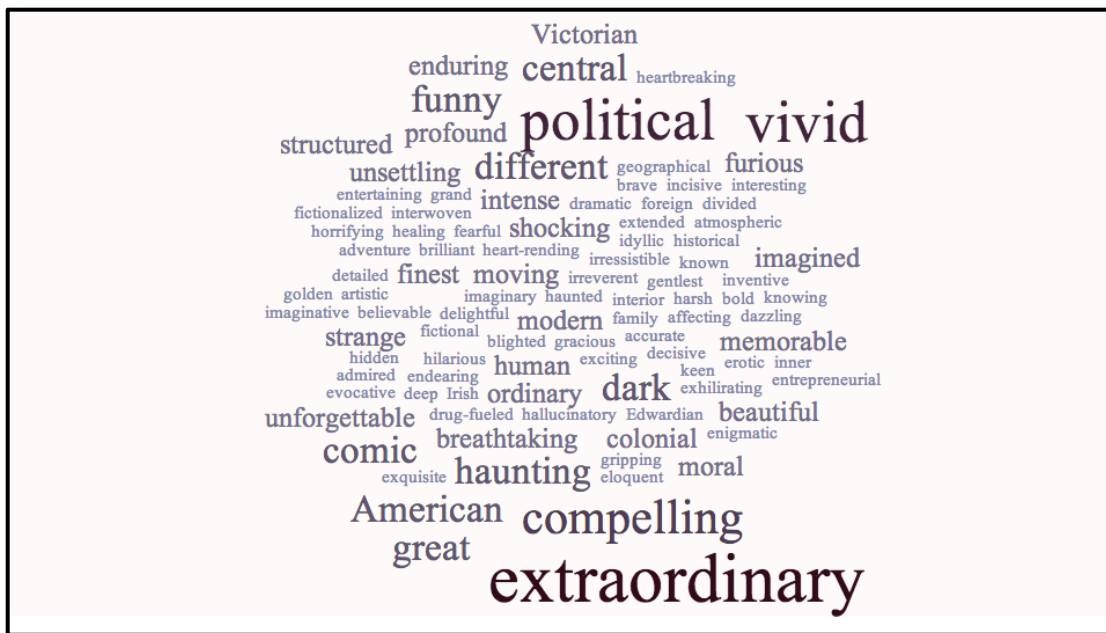
To expand on the individual categories, I also created word clouds for each category. A word cloud takes all the adjectives belonging to a category and arranges them in a cluster, with words that are used more frequently displayed in a larger font than words used less frequently – the larger the word, the more frequently it was used. However, while word clouds ideally “reveal the essential” (Hein), they are not without their failures. Most importantly, word clouds are without context – while we can determine which words are used most frequently, we don’t know how those words were used, the tone with which they were used, or the concept or object to which they were applied. As Graham, Milligan, and Weingart put it, “As adjectives are separated from other concepts, we lose the ability to derive meaning”; that is, in the case of this project, looking at adjectives without the nouns they modify removes a great deal of the meaning from the adjectives. We can gain an understanding of the overall mood or tone of the blurb, but not the objects, people, or situations exemplified by the adjectives creating that mood. Additionally, as Harris claims, they “support only the crudest sorts of textual analysis” – for example, in a word cloud made from a news story about the Iraq War, the words “car” and “blast” may be used frequently, but is the cloud describing exploding vehicles or simply cars and explosions separately? If the cloud reveals that words like “IED” were used frequently, does this tell us anything useful? The conclusions these word clouds allow a reader to reach may be inaccurate or imprecise.

Fortunately, because my word clouds represent specific, tailored groups of words – adjectives that have already been placed into a particular category – they are less susceptible to some of these failures. They do not contain groups of adjectives, nouns, and verbs that fit together in unspecified, inconclusive ways; they don’t contain

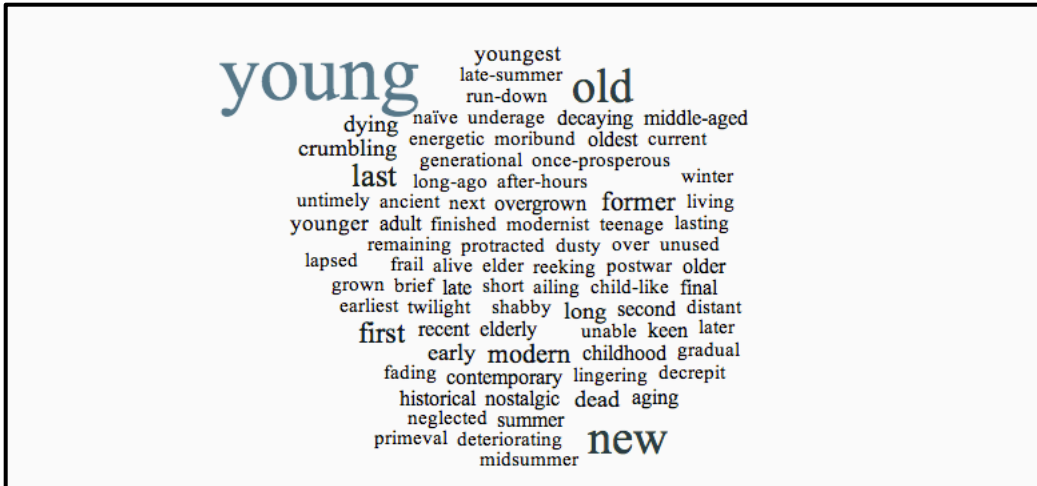


extremely common stopwords like “and” or “don’t” or “the” or “like,” which would throw off the relative sizing of the words; and their membership in specific adjective categories gives them a degree of context within the blurbs that the words in a word cloud depicting all the words in a blurb would not have. While still not perfect, my word clouds can still serve as a useful way to visualize the kinds of words that define and exemplify the blurbs.

Below, I’ve included both the word clouds for the most common categories and the word clouds that have the most telling results – some word clouds, like the word clouds for the masculine and feminine category and the difference and sameness category, essentially depicted two words of approximately the same size each and were not terribly revealing (“male” and “female” and “different” and “same,” respectively).



*Meta Adjectives*



*Aging and Time*



*Violence, Danger, and Brutality*



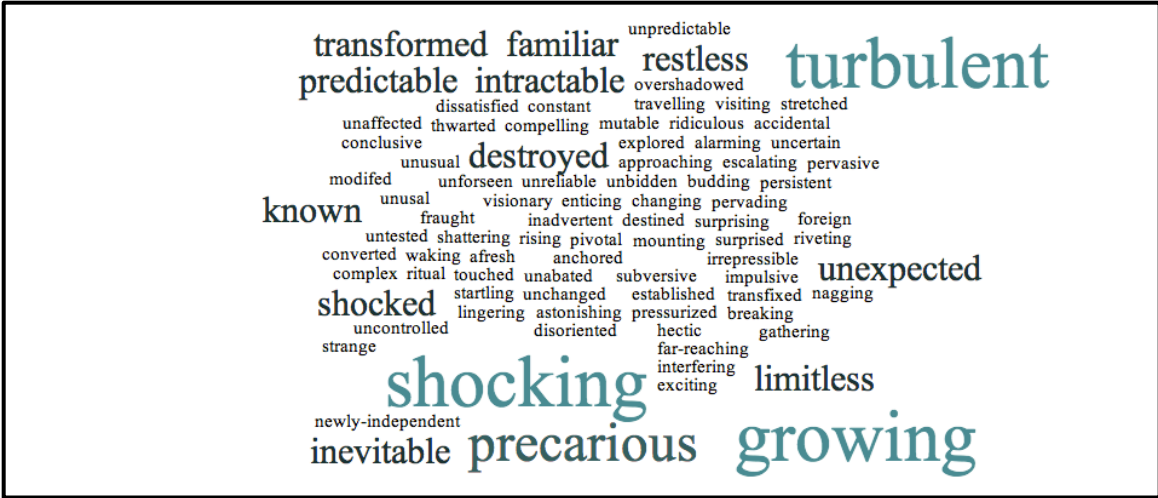
*Nationality and Colonization*



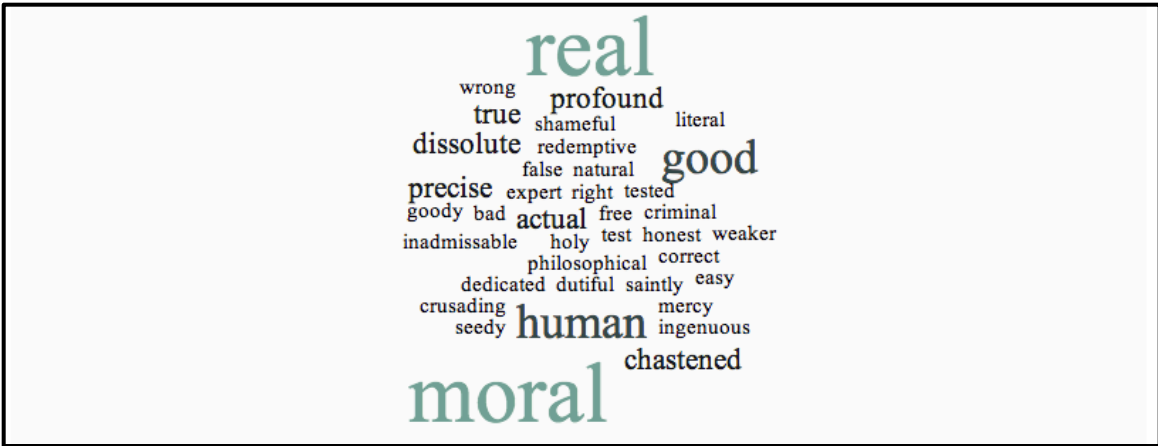
*Class, Money, and Social Reputation*



*Isolation and Interiority*



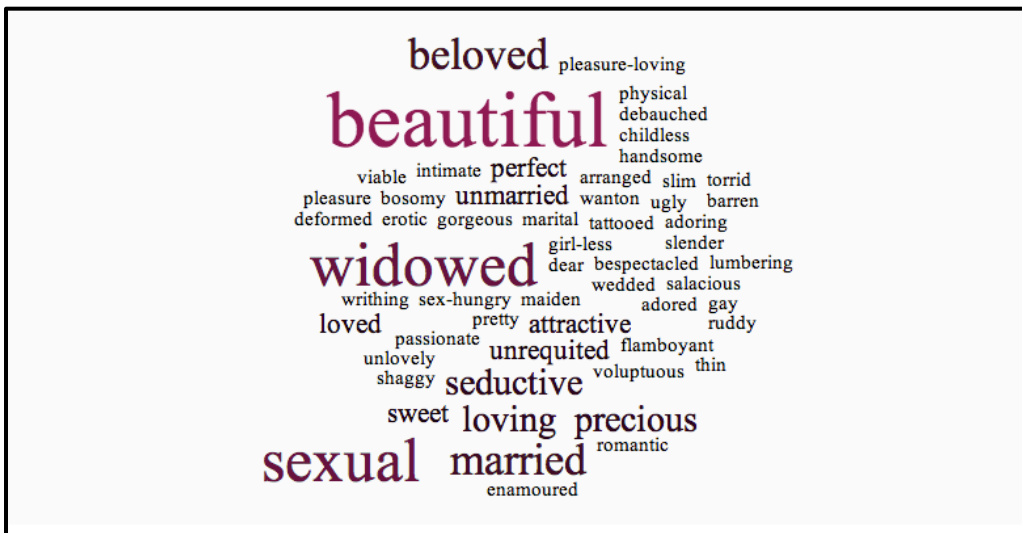
*Permanence and Change*



*Ethics, Truth, and Morality*



*Sadness and Tragedy*



*Sex, Lust, and Beauty*

**family**  
 adoptive intertwined  
**shared** entwined  
 social egalitarian multi-ethnic  
 loyal **close-knit** homemade  
 crowded outside equal  
 linked devoted **fellow**  
 like-minded balanced  
 interracial together cramped  
 peopled utopian  
**diverse** collective  
**entangled**

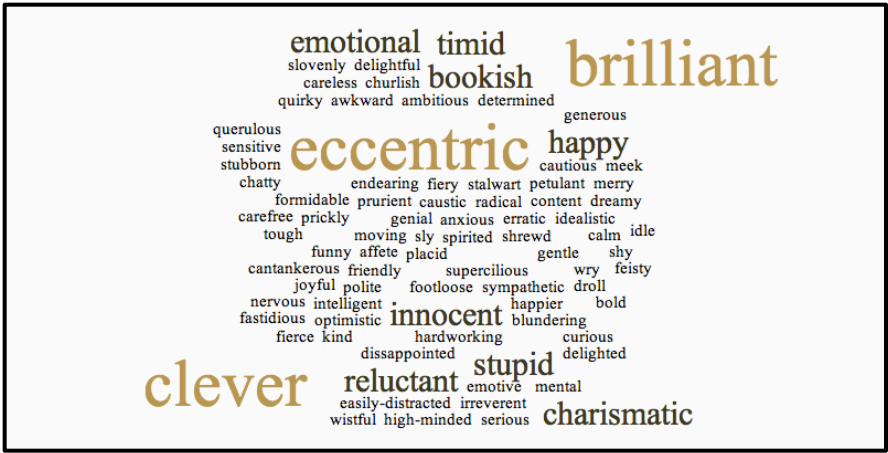
*Community and Kinship*

**mysterious**  
 uncanny **unnamed** inexplicable  
 distorted  
 unsolved enigmatic ritual  
**natural** secret labyrinthine mystical  
 unimaginable strangest  
 spectral unknown bizarre imaginary weird  
 missing experimental folk cryptic **peculiar**  
**elusive** aware ghostly disappearing  
**magical** murky unfamiliar **spiritual**  
 vague odder bewitched  
 unsettling suspicious **unaware** ill-defined  
 supernatural improbable **secretive** ceremonial  
 imagined make-believe telepathic absurd  
**impossible**  
**unlikely** **strange**  
 questionable  
 little-reported

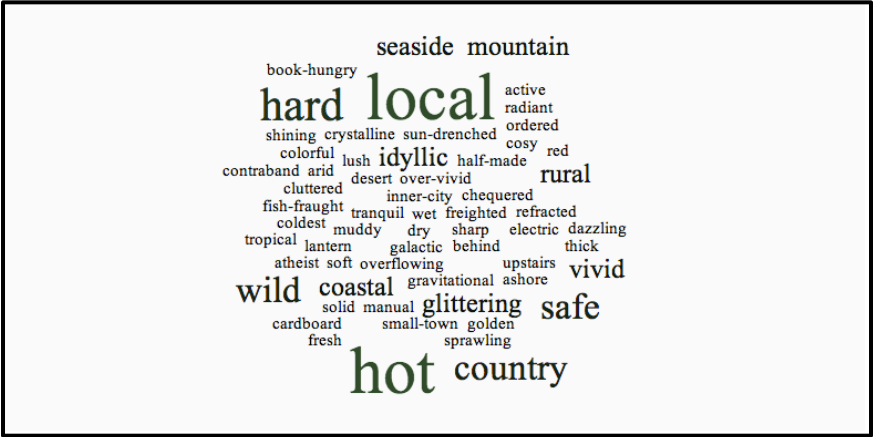
*Mystery*

**great**  
**greatest** **ordinary**  
 unspectacular initial  
 major **epic** electrifying  
**special** **grand** utmost crucial miraculous  
 exemplary favourite ultimate  
 essential significant largest **rare**  
 minor **chief** **outrageous** everyday  
**unusual** legendary leading  
**glorious** intense  
 wonderful spectacular mundane  
 monumental **poetic** remarkable  
 unconventional  
**extraordinary** important  
**powerful**

*Ordinary and Extraordinary*



*Misc. Personality Traits*



*Location*

Examining these word clouds, it appears that the typical Booker Prize novel largely concerns young, individual, and often isolated men, many of whom are fragile, broken, and emotionally damaged. These young men, who are frequently wealthy and respected and most often British or English, interact with the material, violent world around them, coming to understand its darkness and immorality. They encounter turbulent, shocking, and often political events that take their worlds by surprise, and they deal with these events through their brilliance and cleverness. Their experiences with sex and women are frequently in terms of marriage, as evidenced by the recurrence of words like “married” and “widowed,” and these women are typically physically beautiful. If nature or landscape is involved, it is wild and mountainous and mysterious. Ultimately, these stories are not stories about interconnectedness or togetherness or kinship, but about the ordinary individual against the extraordinary events of the world – man versus nature and man versus himself.

Additionally, the keyword analysis provides an interesting way to enter into the three individual texts – specifically, how do the three texts display or not display characteristics of the most frequently occurring or most important categories, and how do these texts illuminate the ways in which the categories might be deployed in an actual novel? This will allow me to gauge whether my keywords analysis has proved to be a sensitive barometer to the characteristics of actual texts and to assess whether the typified Booker Prize novel remains typical when the model is compared with actual nominees.

Looking at the most common category first, it is clear that all three novels rely heavily on themes of aging and time, which, in these texts, are intricately tied to themes of permanence and change. In *The Remains of the Day* and *Cat's Eye*, a single

protagonist reflects back on his or her youth and past, attempting to understand how the present came to be and how he or she fits into the present. Stevens attempts to reconcile his notions of loyalty and duty with modern notions of informality and reduced hierarchy, and attempts to resolve his relationship with Miss Kenton; Elaine attempts to work through the ghosts of her past and attempts to come to terms with the trauma she suffered as a girl, her identity as a feminist painter, and her relationship with Cordelia. In both cases, the temporal distance the narrator has from the past enables the story to function, and the aging process provides a great deal of the story's tension as older characters struggle to come to terms with age and let go of the past.

*On Chesil Beach* does not rely as much on an older character reflecting on youth, though it does incorporate flashbacks and flashforwards from the honeymoon on Chesil Beach, but still relies heavily on unusual temporal tension; the entire novel hinges on a single night and, more specifically, a few drawn out minutes in the bedroom as Florence panics while attempting to have sex with Edward. Entire lives are changed by a singular moment, and all the years before and after that moment that the novel discusses and describes pale in comparison – as Edward puts it, “This is how the entire course of a life can be changed – by doing nothing” (McEwan 203). The past, as in *Cat's Eye* and *Remains*, will haunt the characters for the rest of their lives – as Elaine says in the opening of *Cat's Eye*, “Time is not a line but a dimension, like the dimensions of space” (Atwood 3), and this is how time functions in all three individual texts. Time functions nonlinearly to create tension between the vivid past and the haunted, anguished present or the vivid moment and the haunted, anguished future it will create.



Nationality is another extremely common category that is deployed heavily in all three novels. As we have already seen, the idea of Englishness comes up again and again in both *On Chesil Beach* and *The Remains of the Day*; Englishness is tied to repression and isolation, to outdated tradition and a fading way of life. Stevens cannot escape notions of duty and his identity as, in his American employer's words, a "real old English butler" (Ishiguro 124) – Englishness, for Stevens, is a set of hierarchical, class-based rules and conventions that he cannot follow while living functionally in the modern world. For Edward and Florence, Englishness is the force that prevents them from understanding how to communicate and negotiate their sexuality, from telling each other what they do or do not want. Englishness is crystallized but not idealized – nationality is a source of internal, quiet, male anguish, an inescapable force responsible for a good deal of the characters' actions and flaws.

In *Cat's Eye*, nationality is also a source of quiet anguish, but a much different kind of anguish. For example, one of the paintings at Elaine's retrospective is entitled *The Three Muses* and depicts three figures from her childhood who were in some kind of exile: Mrs. Finestein, her Jewish neighbor during WWII with "death camp ashes blow[ing] daily through [her] head" (Atwood 445); Mr. Banerji, her father's Indian colleague who left Canada when his race kept him from promotion and success; and Miss Stuart, her Scottish school teacher cast "from plundered Scotland still declining, three thousand miles away" (Atwood 445). Here, nationality is not an internalized, crystallized force that controls one's own actions but an external force that allows characters to discriminate against others. In a narrative universe where Canadian national identity is the norm, Atwood does not focus on Elaine's Canadian identity but on the non-Canadian

identities around her. Elaine, a female character, looks outward with empathy, where Edward and Stevens, male characters, look inward with self-loathing. However, in both cases, nationality is not a source of pride but a source of pain – it keeps people from connection and keeps them apart from each other and from their happiness.

The third category that features prominently and importantly in all three novels is isolation and interiority – the protagonists are all, to some extent, isolated from other people and completely focused on their own internal monologues. Stevens can barely relate to other human beings – he interacts with everyone, including Miss Kenton, in a closed-off, unbearably professional way, so that no personal conversation or intimate connection is possible. Edward and Florence’s marriage is destroyed by their own extreme individuality – they do not know how to share themselves with one another, to talk to one another about their insecurities, for fear of proving a disappointment or for fear of humiliation, and so they remain silent and tortured, pushing each other away even as they desperately want to connect. Elaine is the least isolated – she maintains friendships and sexual relationships throughout *Cat’s Eye*, some of which are functional. However, even Elaine struggles to connect to the women around her – she cannot figure out how to perform femininity or feminism in such a way that would allow her to fit in, and as a child, she puts up with terrible torture at the hands of other girls for fear of standing up for herself and losing her only friends. In all these novels, a great deal of tension is derived from the desire to connect and the subsequent failure to connect – loneliness, isolation, and interiority are all causes of anguish as well, and none of the characters seem to really understand how to overcome them.

Of course, this is not an exhaustive list of the categories deployed in the three texts. All three also deal with truth and morality in some parts, and beauty in others, and class in others – it's simply that these categories do not drive and define the texts in the same fundamental way that time, nationality, and isolation do, and they do not seem to be such a source of anguish. It's interesting, however, that just as the more common categories proved most prevalent in the texts, community and kinship, a category that proved uncommon in my keyword analysis, also proved a rarer theme in the individual texts. While the characters in each text have relationships with other people and while they each have family members, there is no sense of community – no sense of group affection, no groups of supportive friends or close-knit towns or large loving families. *Cat's Eye* comes the closest, since Elaine and her brother are initially fairly close with their parents and since Elaine's girlhood group of friends is inseparable, even if they are frequently horrible to each other, but there is still no sense of the redemptive or supportive power of community or friendship or partnership. Even Elaine's ultimate romantic partner, Ben, is primarily interesting to her because he is boring and the antithesis of the dramatic relationships she'd had previously – no one is saved or protected by their love for other people or by other people who love them. Not only are the characters isolated, but they never truly overcome their isolation – they don't find partners or communities or friends who take away the pain that their Englishness or their loneliness or their old age as has brought them.

## Conclusion

Before concluding, it is important to acknowledge the limitations of both distant reading as an approach and this study in particular. For one thing, as Alan Liu points out, “the interpretive or analytical methods at the two ends of the scale, macro and micro, are anything *but* seamless in their relationship” (27); that is, there is nothing between the microscopic focus of a close reading and the macroscopic focus of a project like Moretti’s in *Graphs Maps and Trees*. The broad conclusions drawn from a distant reading may not correspond or even dialog with the deep conclusions drawn from a close one, and there is no midway unit of analysis to bridge the gap. I’ve attempted to deal with this problem by examining how the model of a best text created by my distant reading aligns with closer readings done on individual texts but, as I’ve discovered, while my model aligns well with some individual texts, it doesn’t align with all of them. *Cat’s Eye*, especially, resists the model in many ways – its protagonist is neither male nor English, for example, and the novel is not short. Margaret Atwood is a woman – her woman’s voice is lost in the male authorial perspective my model produces. By nature, dealing with 254 novels at once will give a broad understanding of the novels at hand while excluding the deeper details and anomalies that make some of these novels different and exemplary. Some very long books win the Booker Prize; some are historical fiction, some are science fiction, some are about community and togetherness and are set in unusual times and places.

Additionally, as Amir Khadem points out, a study like the one I’ve just done uses scientific methodology without scientific control – that is, I’m correlating variables without controlling for confounding variables and I’m hypothesizing at causation when

correlation can never equal causation. That is, “the idea of connecting several distinct observations upon an already chosen model borrowed from a scientific field cannot guarantee any significant result” (413), because with literary data, establishing causation to a statistical or scientific standard is not possible. Distant reading attempts to draw scientific conclusions using non-experimental, non-empirical, un-controlled data, and while these conclusions may be illuminating, interesting, and important, they cannot have the degree of certainty we might want them to.

Having acknowledged these limitations, however, this study has allowed me to build a model of the kind of book that is contemporarily considered a best text, using the Booker Prize as a barometer. It has become clear that the Booker Prize does not recognize simply the best novel written in English – rather, the typical Booker Prize nominee fits a distinct model of a contemporary best text. The corpus of shortlisted and winning novels is overflowing with serious, solemn books written by British or Irish men about British or Irish men, books that resist association with genre novels like fantasy or science fiction and books that anchor themselves in reality, in identifiable cities and identifiable countries and identifiable time periods. These books tell the stories of young men who struggle alone through a violent and painful world, who go to war and live in the United Kingdom and participate in the formation and continuation of Englishness and British national identity. Ultimately, it is this gendered, nationalized, isolated pain, set against a backdrop of realism, impending age, and creeping modernity, that pervades the Booker Prize corpus and defines the kind of text that, when using the Booker Prize to sort works that are considered literarily valuable from works that are not, constitutes a member of a contemporary literary canon.

## Works Cited

### *Part I: The Novels*

- Achebe, Chinua. *Anthills of the Savannah*. New York: Anchor, 1988. Print.
- Ackroyd, Peter. *Chatterton*. New York: Grove, 1988. Print.
- Adiga, Aravind. *The White Tiger: A Novel*. New York: Free, 2008. Print.
- Ali, Monica. *Brick Lane: A Novel*. New York: Scribner, 2003. Print.
- Amis, Kingsley. *Ending Up*. New York: Harcourt Brace Jovanovich, 1974. Print.
- Amis, Kingsley. *Jake's Thing*. New York: Viking, 1979. Print.
- Amis, Kingsley. *The Old Devils: A Novel*. New York: Summit, 1987. Print.
- Amis, Martin. *Time's Arrow, Or, The Nature of the Offense*. New York: Harmony, 1991. Print.
- Arden, John. *Silence among the Weapons*. London: Methuen, 1982. Print.
- Atwood, Margaret. *Alias Grace*. New York: Nan A. Talese, 1996. Print.
- Atwood, Margaret. *The Blind Assassin*. New York: N.A. Talese, 2000. Print.
- Atwood, Margaret. *Cat's Eye*. New York: Doubleday, 1988. Print.
- Atwood, Margaret. *The Handmaid's Tale*. Boston: Houghton Mifflin, 1986. Print.
- Atwood, Margaret. *Oryx and Crake*. New York: Nan A. Talese, 2003. Print.
- Azzopardi, Trezza. *The Hiding Place*. New York: Grove, 2002. Print.
- Bailey, Paul. *Gabriel's Lament*. New York: Viking, 1987. Print.
- Bailey, Paul. *Peter Smart's Confessions*. London: J. Cape, 1977. Print.
- Bainbridge, Beryl. *An Awfully Big Adventure*. New York: Harper Collins, 1991. Print.
- Bainbridge, Beryl. *The Bottle Factory Outing*. New York: G. Braziller, 1975. Print.
- Bainbridge, Beryl. *The Dressmaker*. London: Duckworth, 1973. Print.

Bainbridge, Beryl. *Every Man for Himself*. New York: Carroll & Graf, 1996. Print.

Bainbridge, Beryl. *Master Georgie*. New York: Carroll & Graf, 1998. Print.

Ballard, J. G. *Empire of the Sun*. New York: Simon and Schuster, 1984. Print.

Banville, John. *The Book of Evidence*. New York: Scribner, 1989. Print.

Banville, John. *The Sea*. New York: Alfred A. Knopf, 2005. Print.

Barker, A. L. *John Brown's Body*. London: Hogarth, 1969. Print.

Barker, Nicola. *Darkmans*. New York: Harper Perennial, 2007. Print.

Barker, Pat. *The Ghost Road*. New York: Dutton, 1996. Print.

Barnes, Julian. *Arthur & George*. New York: Alfred A. Knopf, 2006. Print.

Barnes, Julian. *England, England*. New York: Alfred A. Knopf, 1999. Print.

Barnes, Julian. *Flaubert's Parrot*. New York: Alfred A. Knopf, 1985. Print.

Barnes, Julian. *The Sense of an Ending*. New York: Alfred A. Knopf, 2011. Print.

Barry, Sebastian. *A Long Long Way*. New York: Viking, 2005. Print.

Barry, Sebastian. *The Secret Scripture*. New York: Viking, 2008. Print.

Bawden, Nina. *Circles of Deceit*. New York: St. Martin's, 1987. Print.

Bedford, Sybille. *Jigsaw*. New York: Alfred A. Knopf, 1989. Print.

Berger, John. *G*. New York: Pantheon, 1980. Print.

Birch, Carol. *Jamrach's Menagerie*. New York: Doubleday, 2011. Print.

Blackwood, Caroline. *Great Granny Webster*. New York: Scribner, 1977. Print.

Booth, Martin. *The Industry of Souls*. New York: Thomas Dunne, 1999. Print.

Bowen, Elizabeth. *Eva Trout; Or, Changing Scenes*. New York: Alfred A. Knopf, 1968.  
Print.

Boyd, William. *An Ice-cream War*. New York: Morrow, 1983. Print.

Bradbury, Malcolm. *Rates of Exchange*. New York: Alfred A. Knopf, 1983. Print.

Brink, André P. *An Instant in the Wind*. New York: Morrow, 1977. Print.

Brink, André P. *Rumours of Rain*. New York: Morrow, 1978. Print.

Brookner, Anita. *Hotel Du Lac*. New York: Pantheon, 1984. Print.

Brown, George Mackay. *Beside the Ocean of Time*. London: J. Murray, 1994. Print.

Burgess, Anthony. *Earthly Powers*. New York: Simon and Schuster, 1980. Print.

Byatt, A. S. *The Children's Book*. New York: Alfred A. Knopf, 2009. Print.

Byatt, A. S. *Possession: A Romance*. New York: Random House, 1990. Print.

Carey, Peter. *Illywhacker*. New York: Harper & Row, 1985. Print.

Carey, Peter. *Oscar & Lucinda*. New York: Harper & Row, 1988. Print.

Carey, Peter. *Parrot and Olivier in America*. New York: Alfred A. Knopf, 2010. Print.

Carey, Peter. *True History of the Kelly Gang*. New York: Alfred A. Knopf, 2000. Print.

Carr, J. L. *The Battle of Pollocks Crossing*. New York: Viking, 1985. Print.

Carr, J. L. *A Month in the Country*. New York: St. Martin's, 1980. Print.

Cartwright, Justin. *In Every Face I Meet*. London: Sceptre, 1995. Print.

Chatwin, Bruce. *Utz*. New York: Viking, 1989. Print.

Coetzee, J. M. *Disgrace*. New York: Viking, 1999. Print.

Coetzee, J. M. *Life & Times of Michael K*. New York: Viking, 1984. Print.

Coetzee, J. M. *Summertime*. Thorndike, ME: Center Point Pub., 2010. Print.

Collins, Michael. *The Keepers of Truth*. New York: Scribner, 2001. Print.

Crace, Jim. *Quarantine*. New York: Farrar, Straus and Giroux, 1997. Print.

Dangor, Achmat. *Bitter Fruit*. New York: Black Cat, 2005. Print.

Davies, Robertson. *What's Bred in the Bone*. New York: Viking, 1985. Print.



Deane, Seamus. *Reading in the Dark*. New York: Alfred A. Knopf, 1997. Print.

Desai, Anita. *Clear Light of Day*. New York: Harper & Row, 1980. Print.

Desai, Anita. *Fasting, Feasting*. New York: Houghton Mifflin, 2000. Print.

Desai, Anita. *In Custody*. New York: Harper & Row, 1984. Print.

Desai, Kiran. *The Inheritance of Loss*. New York: Atlantic Monthly, 2006. Print.

DeWitt, Patrick. *The Sisters Brothers*. New York: Ecco, 2011. Print.

Donoghue, Emma. *Room*. New York: Little Brown and Co., 2010. Print.

Doyle, Roddy. *Paddy Clark Ha Ha Ha*. London: Minerva, 1995. Print.

Doyle, Roddy. *The Van*. New York: Viking, 1992. Print.

Durrell, Lawrence. *Constance, Or, Solitary Practice*. Toronto: General Paperbacks, 1983.  
Print.

Edugyan, Esi. *Half Blood Blues*. London: Serpent's Tail, 2011. Print.

Ellis, Alice Thomas. *The 27th Kingdom*. Wakefield, RI: Moyer Bell, 1999. Print.

England, Barry. *Figures in a Landscape*. New York: Random House, 1968. Print.

Enright, Anne. *The Gathering*. New York: Black Cat, 2007. Print.

Farrell, J. G. *The Siege of Krishnapur*. New York: Harcourt Brace Jovanovich, 1974.  
Print.

Fischer, Tibor. *Under the Frog*. London: Penguin, 1993. Print.

Fitzgerald, Penelope. *The Beginning of Spring*. New York: H. Holt, 1989. Print.

Fitzgerald, Penelope. *The Bookshop*. Boston: Houghton Mifflin, 1997. Print.

Fitzgerald, Penelope. *The Gate of Angels*. New York: Doubleday, 1990. Print.

Fitzgerald, Penelope. *Offshore*. New York: H. Holt, 1987. Print.

Foulds, Adam. *The Quickening Maze*. New York: Penguin, 2010. Print.

Frayn, Michael. *Headlong*. New York: Metropolitan, 1999. Print.

Fuller, John. *Flying to Nowhere: A Tale*. New York: G. Braziller, 1984. Print.

Galgut, Damon. *The Good Doctor*. New York: Grove, 2003. Print.

Galgut, Damon. *In a Strange Room*. London: Atlantic, 2010. Print.

Gardam, Jane. *God on the Rocks*. New York: Europa Editions, 2010. Print.

Ghosh, Amitav. *Sea of Poppies*. New York: Farrar, Straus and Giroux, 2008. Print.

Golding, William. *Rites of Passage*. New York: Farrar, Straus, Giroux, 1980. Print.

Gordimer, Nadine. *The Conservationist*. New York: Viking, 1975. Print.

Grant, Linda. *The Clothes on Their Backs*. New York: Scribner, 2008. Print.

Grenville, Kate. *The Secret River*. New York: Canongate, 2005. Print.

Gunesekera, Romesh. *Reef*. New York: New, 1994. Print.

Gurnah, Abdulrazak. *Paradise*. New York: New, 1994. Print.

Hall, Sarah. *The Electric Michelangelo*. New York: Perennial, 2005. Print.

Hamid, Mohsin. *The Reluctant Fundamentalist*. Orlando, FL: Harcourt, 2007. Print.

Heller, Zoe. *Notes on a Scandal*. London: Viking, 2003. Print.

Hensher, Philip. *The Northern Clemency*. New York: Alfred A. Knopf, 2008. Print.

Hill, Susan. *The Bird of Night*. New York: Saturday Review, 1973. Print.

Hollinghurst, Alan. *The Folding Star*. New York: Pantheon, 1994. Print.

Hollinghurst, Alan. *The Line of Beauty*. New York: Bloomsbury, 2004. Print.

Hope, Christopher. *Serenity House*. London: Macmillan, 1992. Print.

Hulme, Keri. *The Bone People*. Baton Rouge: Louisiana State University Press, 1985.  
 Print.

Hutchinson, R. C. *Rising*. London: Michael Joseph, 1976. Print.

Hyland, M. J. *Carry Me Down*. New York: Canongate, 2006. Print.

Ignatieff, Michael. *Scar Tissue*. New York: Farrar, Straus, and Giroux, 1994. Print.

Ishiguro, Kazuo. *An Artist of the Floating World*. New York: Vintage, 1989. Print.

Ishiguro, Kazuo. *Never Let Me Go*. New York: Alfred A. Knopf, 2005. Print.

Ishiguro, Kazuo. *The Remains of the Day*. New York: Alfred A. Knopf, 1989. Print.

Ishiguro, Kazuo. *When We Were Orphans*. New York: Alfred A. Knopf, 2000. Print.

Jackson, Mick. *The Underground Man*. New York: William Morrow, 1997. Print.

Jacobson, Howard. *The Finkler Question*. London: Bloomsbury, 2010. Print.

Jhabvala, Ruth Praver. *Heat and Dust*. New York: Harper & Row, 1976. Print.

Johnston, Jennifer. *Shadows on Our Skin*. New York: Doubleday, 1978. Print.

Jones, Lloyd. *Mister Pip*. New York: Dial, 2007. Print.

Keane, Molly. *Good Behaviour*. New York: Alfred A. Knopf, 1981. Print.

Kelman, James. *A Disaffection*. London: Secker & Warburg, 1989. Print.

Kelman, James. *How Late It Was, How Late*. New York: W.W. Norton &, 1995. Print.

Kelman, Stephen. *Pigeon English*. London: Bloomsbury, 2011. Print.

Keneally, Thomas. *The Chant of Jimmie Blacksmith*. New York: Viking, 1972. Print.

Keneally, Thomas. *Confederates*. New York: Harper & Row, 1980. Print.

Keneally, Thomas. *Gossip from the Forest*. New York: Harcourt Brace Jovanovich, 1976.  
Print.

Keneally, Thomas. *Schindler's Ark*. London: Coronet, 1983. Print.

Kilroy, Thomas. *The Big Chapel*. London: Faber and Faber, 1971. Print.

Kneale, Matthew. *English Passengers*. New York: Nan A. Talese, 2000. Print.

Lessing, Doris May. *Briefing for a Descent into Hell*. New York: Alfred A. Knopf, 1971. Print.

Lessing, Doris May. *The Good Terrorist*. New York: Alfred A. Knopf, 1985. Print.

Lessing, Doris May. *The Sirian Experiments: The Report by Ambien II, of the Five*. New York: Alfred A. Knopf, 1981. Print.

Levy, Andrea. *The Long Song*. New York: Farrar, Straus and Giroux, 2010. Print.

Levy, Deborah. *Swimming Home*. New York: Bloomsbury, 2012. Print.

Lively, Penelope. *According to Mark*. New York: Beaufort, 1984. Print.

Lively, Penelope. *Moon Tiger*. New York: Grove, 1988. Print.

Lively, Penelope. *The Road to Lichtfield*. London: Penguin, 1983. Print.

Lodge, David. *Nice Work*. New York: Viking, 1989. Print.

Lodge, David. *Small World: An Academic Romance*. New York: Macmillan, 1984. Print.

Mackay, Shena. *The Orchard on Fire*. Wakefield, RI: Moyer Bell, 1996. Print.

MacLaverty, Bernard. *Grace Notes*. New York: W.W. Norton, 1997. Print.

Malouf, David. *Remembering Babylon*. London: Vintage, 1994. Print.

Mantel, Hilary. *Bring up the Bodies*. New York: Henry Holt and Co., 2012. Print.

Mantel, Hilary. *Wolf Hall*. New York: Henry Holt and Co., 2009. Print.

Martel, Yann. *Life of Pi*. New York: Harcourt, 2001. Print.

Mason, Anita. *The Illusionist*. New York: Holt, Rinehart and Winston, 1984. Print.

Matar, Hisham. *In the Country of Men*. New York: Dial, 2007. Print.

Mavor, Elizabeth. *A Green Equinox*. London: M. Joseph, 1973. Print.

Mawer, Simon. *The Glass Room*. New York: Other, 2009. Print.

McCabe, Patrick. *Breakfast on Pluto*. New York: Harper Flamingo, 1998. Print.

McCabe, Patrick. *The Butcher Boy*. New York: Fromm International Pub., 1993. Print.

McCarthy, Tom. *C*. New York: Alfred A. Knopf, 2010. Print.

McEwan, Ian. *Amsterdam*. New York: N.A. Talese, 1999. Print.

McEwan, Ian. *Atonement: A Novel*. New York: Doubleday, 2002. Print.

McEwan, Ian. *Black Dogs*. New York: Nan A. Talese, 1992. Print.

McEwan, Ian. *The Comfort of Strangers*. New York: Simon and Schuster, 1981. Print.

McEwan, Ian. *On Chesil Beach*. New York: Doubleday, 2007. Print.

McGahern, John. *Amongst Women*. New York: Viking, 1990. Print.

Middleton, Stanley. *Holiday*. London: Hutchinson, 1974. Print.

Miller, A. D. *Snowdrops*. London: Atlantic, 2011. Print.

Miller, Andrew. *Oxygen*. New York: Harcourt, 2001. Print.

Mills, Magnus. *The Restraint of Beasts*. New York: Arcade Pub., 1998. Print.

Mistry, Rohinton. *Family Matters*. New York: Alfred A. Knopf, 2002. Print.

Mistry, Rohinton. *A Fine Balance*. New York: Alfred A. Knopf, 1996. Print.

Mistry, Rohinton. *Such a Long Journey*. New York: Alfred A. Knopf, 1991. Print.

Mitchell, David. *Cloud Atlas*. New York: Random House, 2004. Print.

Mitchell, David. *number9dream*. New York: Random House, 2001. Print.

Mo, Timothy. *An Insular Possession*. New York: Random House, 1987. Print.

Mo, Timothy. *The Redundancy of Courage*. London: Chatto & Windus, 1991. Print.

Mo, Timothy. *Sour Sweet*. New York: Vintage, 1985. Print.

Moore, Alison. *The Lighthouse*. London: Salt, 2012. Print.

Moore, Brian. *The Colour of Blood*. London: Cape, 1987. Print.

Moore, Brian. *The Doctor's Wife*. New York: Farrar, Straus and Giroux, 1976. Print.

Moore, Brian. *Lies of Silence*. New York: Doubleday, 1990. Print.

Morrall, Clare. *Astonishing Splashes of Colour*. New York: Harper Collins, 2004. Print.

Morris, Jan. *Last Letters from Hav*. New York: Random House, 1985. Print.

Mosley, Nicholas. *Impossible Object*. New York: Coward-McCann, 1969. Print.

Munro, Alice. *The Beggar Maid*. New York: Alfred A. Knopf, 1979. Print.

Murdoch, Iris. *The Black Prince*. New York: Viking, 1973. Print.

Murdoch, Iris. *The Book and the Brotherhood*. New York: Viking, 1988. Print.

Murdoch, Iris. *Bruno's Dream*. New York: Viking, 1969. Print.

Murdoch, Iris. *The Good Apprentice*. New York: Viking, 1986. Print.

Murdoch, Iris. *The Nice and the Good*. New York: Viking, 1968. Print.

Murdoch, Iris. *The Sea, the Sea*. New York: Viking, 1978. Print.

Naipaul, V. S. *A Bend in the River*. New York: Alfred A. Knopf, 1979. Print.

Naipaul, V. S. *In a Free State*. New York: Alfred A. Knopf, 1971. Print.

Newby, P. H. *Something to Answer For*. Philadelphia: Lippincott, 1969. Print.

O'Doherty, Brian. *The Deposition of Father McGreevy*. New York: Turtle Point, 1999.  
 Print.

O'Faolain, Julia. *No Country for Young Men*. New York: Carroll & Graf, 1986. Print.

O'Hagan, Andrew. *Our Fathers*. New York: Harcourt Brace, 1999. Print.

Okri, Ben. *The Famished Road*. London: Vintage, 1992. Print.

Ondaatje, Michael. *The English Patient*. New York: Alfred A. Knopf, 1992. Print.

Parks, Tim. *Europa*. New York: Arcade Pub., 1998. Print.

Phillips, Caryl. *Crossing the River*. New York: Alfred A. Knopf, 1994. Print.

Pierre, D. B. C. *Vernon God Little*. New York, NY: Canongate, 2003. Print.

Pym, Barbara. *Quartet in Autumn*. New York: Dutton, 1978. Print.

Rathbone, Julian. *Joseph*. London: Abacus, 1999. Print.

Rathbone, Julian. *King Fisher Lives*. New York: St. Martin's, 1976. Print.

Richler, Mordecai. *Solomon Gursky Was Here*. New York: Alfred A. Knopf, 1990. Print.

Richler, Mordecai. *St. Urbain's Horsemen*. Toronto: McClelland & Stewart, 1971. Print.

Roberts, Michele. *Daughters of the House*. New York: W. Morrow, 1992. Print.

Robinson, Derek. *Goshawk Squadron*. New York: Viking, 1972. Print.

Roy, Arundhati. *The God of Small Things*. New York: Random House, 1997. Print.

Rubens, Bernice. *The Elected Member*. London: Abacus, 1986. Print.

Rubens, Bernice. *A Five Year Sentence*. London: W.H. Allen, 1978. Print.

Rushdie, Salman. *Midnight's Children*. New York: Alfred A. Knopf, 1981. Print.

Rushdie, Salman. *The Moor's Last Sigh*. New York: Pantheon, 1995. Print.

Rushdie, Salman. *The Satanic Verses*. New York, NY: Viking, 1989. Print.

Rushdie, Salman. *Shame*. New York: Alfred A. Knopf, 1983. Print.

Schlee, Ann. *Rhine Journey*. New York: Holt, Rinehart and Winston, 1981. Print.

Scott, Paul. *Staying On*. New York: Morrow, 1977. Print.

Seiffert, Rachel. *The Dark Room*. New York: Pantheon, 2001. Print.

Self, Will. *Umbrella*. New York: Grove, 2012. Print.

Shields, Carol. *The Stone Diaries*. New York: Viking, 1994. Print.

Shields, Carol. *Unless*. London: Fourth Estate, 2002. Print.

Sinha, Indra. *Animal's People*. New York: Simon & Schuster, 2008. Print.

Smith, Ali. *The Accidental*. New York: Pantheon, 2005. Print.

Smith, Ali. *Hotel World*. New York: Anchor, 2002. Print.

Smith, Zadie. *On Beauty: A Novel*. New York: Penguin, 2005. Print.

Snow, C. P. *In Their Wisdom*. New York: Scribner, 1974. Print.

Soueif, Ahdaf. *The Map of Love*. New York: Anchor, 2000. Print.

Spark, Muriel. *Loitering with Intent*. New York: Coward, McCann & Geoghegan, 1981. Print.

Spark, Muriel. *The Public Image*. New York: Alfred A. Knopf, 1968. Print.

St Aubyn, Edward. *Mother's Milk*. New York: Open City, 2005. Print.

St John, Madeleine. *The Essence of the Thing*. New York: Carroll & Graf, 1998. Print.

Storey, David. *Pasmore*. New York: Dutton, 1974. Print.

Storey, David. *Saville*. New York: Harper & Row, 1976. Print.

Swift, Graham. *Last Orders*. New York: Alfred A. Knopf, 1996. Print.

Swift, Graham. *Waterland*. New York: Poseidon, 1983. Print.

Tan, Twan Eng. *The Garden of Evening Mists*. Newcastle upon Tyne: Myrmidon, 2012. Print.

Taylor, Elizabeth. *Mrs. Palfrey at the Claremont*. New York: Viking, 1971. Print.

Thayil, Jeet. *Narcopolis*. London: Faber, 2012. Print.

Thomas, D. M. *The White Hotel*. New York: Viking, 1981. Print.

Tóibín, Colm. *The Blackwater Lightship*. New York: Scribner, 2000. Print.

Tóibín, Colm. *The Master*. New York: Scribner, 2004. Print.

Toltz, Steve. *A Fraction of the Whole*. New York: Spiegel & Grau, 2008. Print.

Tremain, Rose. *Restoration*. New York: Viking, 1990. Print.

Trevor, William. *The Children of Dynmouth*. New York: Viking, 1977. Print.

Trevor, William. *Mrs. Eckdorf in O'Neill's Hotel*. London: Bodley Head, 1969. Print.



Trevor, William. *Reading Turgenev*. London: Viking, 1991. Print.

Trevor, William. *The Story of Lucy Gault*. New York: Viking, 2002. Print.

Unsworth, Barry. *Morality Play*. New York: N.A. Talese, 1995. Print.

Unsworth, Barry. *Pascali's Island*. London: Joseph, 1980. Print.

Unsworth, Barry. *Sacred Hunger*. New York: Doubleday, 1992. Print.

Walsh, Jill Paton. *Knowledge of Angels*. Boston: Houghton Mifflin, 1994. Print.

Warner, Marina. *The Lost Father*. New York: Simon and Schuster, 1988. Print.

Waters, Sarah. *Fingersmith*. New York: Riverhead, 2002. Print.

Waters, Sarah. *The Little Stranger*. New York: Riverhead, 2009. Print.

Waters, Sarah. *The Night Watch*. New York: Riverhead, 2006. Print.

Weldon, Fay. *Praxis*. New York: Summit, 1978. Print.

Williams, Gordon M. *From Scenes like These*. New York: Morrow, 1969. Print.

Winton, Tim. *Dirt Music*. New York: Scribner, 2002. Print.

Winton, Tim. *The Riders*. New York: Scribner, 1995. Print.

Woodward, Gerard. *I'll Go to Bed at Noon*. London: Chatto & Windus, 2004. Print.

*Part II: Theoretical and Reference Texts*

“The Book Genome Project.” *The Book Genome Project*. N.p., n.d. Web. 10 April 2014.  
<<http://www.thebooklamp.org/>>.

Franková, Milada. “The (Booker) Prize For Fiction And The Tenor Of The Times.” *Fiction and Literary Prizes in Great Britain*. 53-58. Vienna, Austria: Praesens, 2006. Print.

Graham, Shawn, Ian Milligan, and Scott Weingart. "Basic Text Mining: Word Clouds, Their Limitations, and Moving Beyond Them." *The Historian's Macroscope: Big Digital History*. 11, 2013. Web. 20 April 2014.

<<http://www.themacroscope.org/>>.

Guillory, John. *Cultural capital the problem of literary canon formation*. Chicago: University of Chicago Press, 1993. Print.

Harris, Jacob. "Word Clouds Considered Harmful." *Nieman Journalism Lab*. N.p., 13 October 2011. Web. 20 April 2014.

<<http://www.niemanlab.org/2011/10/word-clouds-considered-harmful/>>.

Hein, Robert. "Word Clouds – Love ‘em or Hate ‘em?" *Vision Critical*. N.p., 9 August 2012. Web. 20 April 2014.

<<http://www.visioncritical.com/blog/pros-and-cons-word-clouds-visualizations>>.

Jockers, Matthew Lee. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois, 2013. Print.

Khadem, Amir. "Annexing the Unread: A Close Reading of 'Distant Reading.'" *Neohelicon* 39.2 (2012): 409-21. Print.

Liu, Alan. "The State of the Digital Humanities: A Report and a Critique." *Arts and Humanities in Higher Education* 11.1-2 (2012): 8-41. Print.

"The Man Booker Prizes." *The Man Booker Prizes*. N.p., n.d. Web. 10 April 2014.

<<http://www.themanbookerprize.com/>>.

Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for a Literary History*. London: Verso, 2005. Print.

- Moseley, Merritt. "Margins Of Fact And Fiction: The Booker Prize 2009." *Sewanee Review* 118.3 (2010): 429-435. Print.
- Moseley, Merritt. "Tidy And Untidy Novels: The Booker Prizes 2006-2007." *Sewanee Review* 116.2 (2008): 289-303. Print.
- Mukherjee, Ankhi. "'What Is A Classic?': International Literary Criticism And The Classic Question." *PMLA: Publications Of The Modern Language Association Of America* 125.4 (2010): 1026. Print.
- Norris, Sharon. "Recontextualising The Booker." *Fiction and Literary Prizes in Great Britain*. 20-36. Vienna, Austria: Praesens, 2006. Print.
- Todd, Richard. *Consuming Fictions: The Booker Prize And Fiction In Britain Today*. London: Bloomsbury, 1998. Print.