

Reader and author gender and genre in Goodreads

Mike Thelwall

University of Wolverhampton, UK

Journal of Librarianship and

Information Science

2019, Vol. 51(2) 403–430

© The Author(s) 2017

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0961000617709061

journals.sagepub.com/home/lis



Abstract

There are known gender differences in book preferences in terms of both genre and author gender but their extent and causes are not well understood. It is unclear whether reader preferences for author genders occur within any or all genres and whether readers evaluate books differently based on author genders within specific genres. This article exploits a major source of informal book reviews, the Goodreads.com website, to assess the influence of reader and author genders on book evaluations within genres. It uses a quantitative analysis of 201,560 books and their reviews, focusing on the top 50 user-specified genres. **The results show strong gender differences in the ratings given by reviewers to books within genres, such as female reviewers rating contemporary romance more highly, with males preferring short stories. For most common book genres, reviewers give higher ratings to books authored by their own gender, confirming that gender bias is not confined to the literary elite.** The main exception is the comic book, for which male reviewers prefer female authors, despite their scarcity. A word frequency analysis suggested that authors wrote, and reviewers valued, gendered aspects of books within a genre. For example, relationships and romance were disproportionately mentioned by women in mystery and fantasy novels. These results show that, perhaps for the first time, **it is possible to get large-scale evidence about the reception of books by typical readers, if they post reviews online.**

Keywords

Book reviews, Goodreads, social web

Introduction

Gender differences in book authorship and book preferences have been known about for a long time, but their causes are not fully understood (see the section below). In particular, some genres are known to be written more and enjoyed more by females, and some by males. Nevertheless, it is not clear whether this partiality also occurs *within* some or all genres and whether the relationship between writer and reader genders shapes *attitudes* towards books within genres amongst those that choose to read them. This knowledge would help to gain new insights into the role of gender in book reading. This information would be particularly useful to publishers and authors that are attempting to find audiences for written works. One way to address these issues is to examine readers' book reviews (Baym, 1987; Milota, 2014).

Gender differences in book preferences may extend to gender differences in book reviewing. The highest profile book reviews are those published in literary magazines and respected newspapers. Literary criticism has a long history of gender bias (Castle, 1997) and books today by male

authors still seem more likely to be selected for review by prestigious periodicals than books by female authors (VIDA, 2015). It is unclear whether gender differences in the selection of books for review are the preserve of the literary elite, professional reviewers in general or the wider reviewing public. It is also unclear whether this bias also extends to judgements about books in reviews.

Book reviews have been written and shared for millennia (Ortega y Miranda, 1996) but the first systematic attempt to publish them in English was the 1749 *Monthly Review* (Monthly Review, 1844: 1). This reviewed a wide range of printed material, including books, magazines and plays. At the time, printing had made possible an explosion in reading and publishing so that increasingly many people needed to make purchasing or borrowing decisions about

Corresponding author:

Mike Thelwall, School of Mathematics and Computing, University of Wolverhampton, Wolverhampton, WV1 1LY, UK.

Email: M.Thelwall@wlv.ac.uk

an expanding range of new titles (Brewer, 2013). The social web has produced a second transformation with websites like Goodreads.com and Amazon.com allowing ordinary readers to publish book reviews.

This article uses data from reviews posted to the Goodreads website for a random selection of over 500,000 books, analysing gender differences separately for the biggest 50 user-defined genres. The goal is to assess how widespread reader/writer gender preferences are amongst general book readers and to use the results to inform a discussion about the potential causes of any preferences found. This seems to be the first large-scale analysis of reader/writer gender preferences in amateur book reading. It extends the results of a previous analysis of Goodreads (Thelwall, 2017) by taking reader gender into account and analysing the content of reviews.

Background

Books can be read for many purposes, from education to recreation. Novels also have different attractions, from escapism to education, learning, polemicizing or cultural enrichment. Whilst some uses of books might be straightforward, such as gardening advice from a reference work, others may be more complex and have gender influences. To give stereotyped examples of escapism, an 80-year-old woman in a nursing home might enjoy the escapism of a category romance about a 21-year-old nurse attracting an older doctor, and a happily married man with vertigo that opposes capital punishment might still become engrossed in a tale of a spy who parachuted in to kill the villain and get the girl. One of the pleasures of escapism is not having to critically evaluate the content (Henning and Vorderer, 2001).

This section discusses book authorship, readership and reviewing as well as book genres and the role of gender within each of these. The focus is on books written in English, unless otherwise specified.

Gender and authorship

Historically, the majority of writers of English language published books were male until about the middle of the 18th century (Prescott, 2003; Turner, 1992). Writing was considered unfeminine and women that broke the norms risked criticism. For example, Margaret Cavendish (1623–1673) included within her books explicit justifications for writing rather than engaging in more feminine activities (Rees, 2003). In contrast, the first English female professional writer, the dramatist and pioneering novelist Aphra Behn (c. 1640–1689), unapologetically shattered gendered norms of propriety (Hutner, 1993; Young, 1993). Her uniqueness may even have been an advantage in the competitive world of Restoration theatre (Gallagher, 1988). There have subsequently been

complex fluctuations in the proportions and fortunes of female authors, including variations by geographic location and cultural identity (Jockers, 2013).

By the end of the 19th century female authors were apparently outselling males (Stanton, 1988), with this rise leading to public complaints from male writers (e.g. Nelson, 2001). A male elite then identified a subset of all published works as ‘literary’, separating them out as having enhanced social value and greater intrinsic worth (Tuchman and Fortin, 2012). The authors of the selected books were overwhelmingly male. Part of the method of establishing dominance was through the *selection* of books to review in prestigious weeklies (Casey, 1996; Harsh, 2001). A tradition then started in which high value, and predominantly male-authored books were accepted into a literary canon to form the foundation of valued fiction in English (Guillory, 2013). This controversial process (e.g. Verboord, 2003) effectively side-lined women from prestigious writing, but this gender difference has shrunk over time. The *New York Times* Adult Fiction Bestsellers List 1960–2015 gives an indication of ongoing gender differences in the most popular books, since a majority were written by solo males (56%) in comparison to 37% by solo females, although the difference seems to be shrinking (Ortiz, 2015). More generally, there seems to be a pattern where the social status of a cultural activity correlates with the extent to which male artists dominate, apparently due to a preponderance of male evaluators (e.g. award committees, journalists, reviewers) (Verboord, 2012).

Although men and women both write all types of books, albeit in differing proportions (see below), there may be a tendency to use different approaches to writing stories. For example, men may tend to develop goal-driven plots (Brooks, 1992) and women more nurturing stories (Winnett, 1990), but there is no systematic evidence of this and there are many exceptions (e.g. Upton, 1991). Irrespective of genre, there may be themes that are addressed more by one gender than another, such as ‘affection and happiness’ and ‘female fashion’ in 19th-century books (Jockers, 2013: Ch. 8).

Gender and book reading

In the developed world, high rates of literacy make the decision to read a book for pleasure a matter of personal taste, although books may also need to be read for work purposes. Since the middle of the 20th century, book reading for pleasure seems to have declined in popularity, probably due to competition from alternative forms of entertainment, such as TV, smartphones and (perhaps) the Internet (e.g. Beutel et al., 2011; Bureau of Labor Statistics, 2015; Griswold et al., 2005; Smith, 2009). Nevertheless, there remains an increasingly small minority of regular book readers, the ‘reading class’ (Griswold et al., 2005) for whom reading books is an important activity. This reading

class is predominantly well educated and with a high proportion of females (Griswold et al., 2005). In fact, women seem to read more books than men in the developed nations for which survey data is available, including Finland, France, Hong Kong, New Zealand, Poland, Russia, Spain (including magazines) and the UK (Cushman et al., 2005: 67, 89, 118, 182, 201, 224, 237; Bureau of Labor Statistics, 2015).

In the USA, women are more proficient readers but read more fiction at least partly due to differing gender socialization and gender roles. Proficient male readers tend instead to read more non-fiction (Tepper, 2000). Fiction reading seems to be socialized during childhood as a feminine activity and accepted as such by adults that conform to traditional gender roles (Tepper, 2000). The evidence for this is survey data from 1982 and 1992, however, and it is likely that gendered socialization patterns have reduced and changed in the 24 years since then. For example, there are now more women than men in higher education (Goldin et al., 2006).

Gender differences in other patterns of cultural consumption, such as television watching and video game playing, are also widespread (Bihagen and Katz-Gerro, 2000; Cherney and London, 2006) and may influence the time available for book reading. Gender differences occur even for adults within the same occupations (e.g. Lizardo, 2006), suggesting that gender is the core explanation rather than a by-product of other structural differences in society. Nevertheless, gender differences are not immutable and some cultural activities, such as attending opera, have experienced gender switches (Söderberg, 1998).

Gender and genre

Despite the existence of widely used categorization schemes for books (e.g. Dewey, Library of Congress) and widely recognized fiction genres, there is no authoritative genre list or set of major definitions. For fiction, genre is a loose concept that often relates to the topic and structure of a novel (Pyrhönen, 2007; Todorov, 1990). Within Goodreads, the concept encompasses common ways of categorizing books by content, including the broad dichotomy fiction/non-fiction as well as topic/structure genres, such as paranormal romance. Goodreads genres also include *classic* for long-term literary value, formats such as comic books, and topic categorizations, such as science, history or animal, that would not fit standard genre conceptions. Over time, genres can gradually evolve, mutate into something different, combine with other genres and split into subgenres. Users, writers and publishers may interpret genres in their own ways, booksellers may invent new genres to market books, libraries may categorize books to attract readers, and genre types and typical structures can vary internationally.

Although there is no consensus about book genres, there are recognized types, even if some are controversial.

A major distinction that relates to the traditional concept of genre is between literature (i.e. 'highbrow' fiction), mainstream fiction and genre/popular fiction, with the latter tending to closely conform to the genre norms of the day. The boundaries between these are subjective and fluid.

Women are the most successful books within some genres, whereas men are more successful in others. For instance, more female-authored children, young adult and romance books were bestsellers in the UK in 2013 as were more male-authored crime, thriller and adventure books (Cherny, 2012).

From the perspective of gender, romantic fiction is the most important genre because of its high gender imbalance. Romantic fiction is overwhelmingly female authored, and has been for over a century (Hipsky, 2011). For instance, in 2008, one male and 199 female authors wrote Mills and Boon category romances (Jackson, 2008), although the proportion of male authors may well be substantially higher for other romances. The category romance genre is often regarded as being of little literary merit (e.g. Tapper, 2014), which may taint perceptions of the value of female-oriented fiction overall. Chick-lit (Ferriss and Young, 2013) is an extreme case of fiction that is ostensibly written exclusively for one gender. In contrast, science fiction is traditionally male dominated, albeit evolving (e.g. Isaacson, 2013; Rashley, 2007). These gender differences in the consumption of fiction genres may echo gendered socializing and social constraints (e.g. Radway, 1984) but the fundamental causes are unclear.

Book reviews

A book review is a critical evaluation of a published book. A review may give a description of the nature of a work, a critical evaluation of its strong and weak points and a final evaluation. Reviews are typically written by someone other than the author in order to be credible, and vary in length from a sentence or short paragraph to a few pages. A summary review just describing the contents of a book may be most suitable for non-fiction and when the target audience is librarians. In contrast, an opinion review may also give the reviewer's opinions about a volume and may be most suitable for fiction. Scholarly reviews are typically more in depth and critical, and are often published in academic journals (Moreno and Suárez, 2008). Reviews that also attempt to interpret a text or evaluate it in more detail may instead be classified as literary criticism, which is the study, evaluation, and interpretation of literature (Charles, 2011).

Professional reviewers may follow standard guidelines for writing particular types of reviews (e.g. children's books: Horning, 2010; library recommendations: Hooper, 2010). It seems likely that a higher proportion of amateur reviews would be less structurally constrained to the extent that they are not exactly reviews. For instance, it would not

be surprising to see an uncritical hymn of praise from a fan of the writer or a single swear word from a disgruntled customer. Despite this, the prevalence of online reviews seems to have undermined the authority of the expert reviewer (Verboord, 2009). Sites like Goodreads that allow amateur reviewers to write about books may also threaten authors because professional reviews tend to be positive (Little, 2014). The willingness of readers to be critical of a book is an advantage from the perspective of analysing reactions to books, however.

The main purpose of a professional book review is presumably to provide a critical evaluation of a book to aid others in deciding whether to buy it. Nevertheless, a reviewer may also wish to display their literary expertise, highlight an issue discussed in the book (e.g. domestic violence), point out inaccuracies (especially in non-fiction), argue against, or in support of, the main thesis in the book, or provide a perspective to help others to evaluate or discuss a book after reading it (Foster, 1951; Wallace, 1987). All of these reasons may also apply to amateur reviewers but they may also wish to write about a book as an exercise for themselves or their friends, to practise their language skills (Osmani, 2015) or they may have apparently more trivial reasons such as to joke about the book by writing a review in the style of the book as a parody (Posner, 1992). Writing an amateur review in a social website can also turn reading into a social act in the same way as for book reading groups (Steiner, 2008). Thus, amateur book reviews may well have a different function and structure to professional book reviews of all types. Of course, some apparently amateur reviewers also write at the behest of the author or publisher to promote a book (Roberts, 2015) or may be incognito authors praising their own work.

Book reviews are commonly published in dedicated periodicals and in sections of academic journals, newspapers and magazines. Popular sources have been criticized for being conservative in their choice of topics for book reviews (Cleland, 1984) and there are national and cultural biases in the selection of books to review in addition to the gender biases discussed above. For example, non-western writers seem to have the best chance of getting their work reviewed when there is a historical connection between their country of origin or a diaspora community in the western nation (Berkers et al., 2011).

Reviews within major newspapers include subjective judgements about the quality of the books reviewed. These quality judgements seem to be fundamentally instinctive or emotional and hence cannot be full objective (Chong, 2013). Reviewers may empathize with the reviewed authors, giving a tendency to tone down negative reviews (Chong, 2015). Professional reviewers may write to showcase their literary abilities or to promote their own perspective, which would be unethical (Xu and Xi, 2013). In addition, they may be influenced by factors other than the quality of the work reviewed, such as publisher status and

author actions (Janssen, 1998; Van Rees, 1987). In order to reduce the chance of introducing biases, reviewers may avoid others' reviews of the same book, provide a rational explanation for their decision, and check their work for fairness (Chong, 2013, 2015).

Published reviews of fiction seem to be predominantly positive, but with reviewers frequently giving different ratings from each other in their evaluations, suggesting that subjective personal judgements are important in the enjoyment of stories (Palmer, 1995). Subjectivity may be particularly evident in reviews of some genres, such as gay and lesbian fiction (Rothbauer and McKechnie, 2000), presumably reflecting social attitudes, religious or political beliefs. Nevertheless, fiction reviewers seem to be cautious due to concerns about the effect on their reputation of publishing an incorrect judgment about a book – and particularly a negative judgement about a novel that becomes well received (Janssen, 1997). Conversely, less cautious reviewers may welcome the chance to make a name for themselves by publishing a minority opinion (Chong, 2015). For literary fiction, critics' opinions seem to lead to a consensus about the quality of a work, although this is socially constructed judgement rather than an empirically verifiable one (Van Rees, 1989).

Book reviews are sometimes useful for library purchasing decisions (Palmer, 1987, 1991), although academic book reviews published in journals are often not timely enough for this (Lauer, 1989; *Choice* is an exception: Wheeler Carlo et al., 1998). Positive customer reviews of books seem to lead to increased online sales but very negative reviews are more influential (Chevalier and Mayzlin, 2006). In particular, potential readers may evaluate the critical points made in negative reviews (Lin et al., 2005). Thus, some customers clearly read reviews prior to making a purchase rather than relying just on ratings (Chevalier and Mayzlin, 2006), and the content of online reviews also influences library borrowing decisions (Kuei Huang and Yang, 2010). The usefulness of an online review is not the same for all product types (Mudambi and Schuff, 2010), and other factors influence book purchases, such as the reputation of the author and publisher, and the book cover (d'Astous et al., 2006).

Gender and reviewing

There have been allegations of continuing gender bias in the reception of literary novels, such as for prestigious book reviews (VIDA, 2015). These allegations presume that similar numbers of men and women aspire to writing literary works, whereas this is unproven. One study from authors starting in 1955 found that a lower percentage of first books by women were reviewed in the *New York Times*, giving some evidence of systematic selection bias over half a century ago, but no evidence of gender bias in judgements within the published reviews (Ekelund and

Börjesson, 2002). Women seem to write the majority of books and so it seems likely that more women in total aspire to write literary novels. This argument pre-supposes that the concept of the literary novel is meaningful but many authors and others would disagree (Guillory, 2013). Given the above hypothesis of male dominance amongst evaluators, it seems likely that reviewing in an environment, such as Goodreads, that is not dominated by males may not show the same gender bias in reviewer judgments. This has been shown in that female authors of new popular fiction in February 2009 in the USA were less likely than males to receive a literary newspaper review (9% against 12%: not statistically significant difference) but were much more likely to be reviewed in Goodreads (81% against 61%) (Verboord, 2011).

A study of professional book reviews found that reviewers tended to be more positive about books written by people of the same gender (Moore, 1978). Within academia, female political scientists wrote an increasing proportion of books and book reviews between 1991 and 2001 (McGinty and Moore, 2008) and males and females now seem to use similar styles in scholarly book reviews (Tse and Hyland, 2009).

Gender in the social web

There are gender differences in patterns of social web use, although these are probably culture specific. In general, women seem to be more active in posting content to social network sites (Thelwall, 2008), express more positive sentiment (Thelwall et al., 2010) and make more friend connections (Thelwall, 2008). Although women are more concerned about privacy and more likely to hide their identity (Tufekci, 2008), this does not seem to inhibit their online interactions (Stutzman et al., 2011). There are also gender differences in the uptake of major social network sites but these are less important than personality differences in the motivations for using them (Ellison et al., 2007; Muscanell and Guadagno, 2012). Especially for younger users, posting content to the Web is a form of identity performance (Boyd, 2007) and they therefore need to consider how others will view their activities, including perhaps whether they are gender normative.

Gender homophily might be expected in the social web since it occurs offline for friendship, although it is not particularly strong in the USA (McPherson et al., 2001). Members of the major (at the time) social network site MySpace in 2008 did not comment more to members of the same gender, even though they did comment more to people with the same religion, ethnicity and age (Thelwall, 2009). The reason could be that young adult users wanted to display connections to people of the opposite gender. In MySpace in 2007 and Facebook in 2011, there was gender homophily for friendship for females but not for males: both preferred female friends (Thelwall, 2008; Ugander

et al., 2011). Thus, adult females may be better communicators or more active networkers overall. In contrast, a study of language use in Twitter found gender homophily in the content of tweets due to the existence of topics dominated by one gender (Bamman et al., 2014).

There are clear gender differences in book-based social websites. Goodreads members are about 76% female (Thelwall and Kousha, 2017) and there are substantial gender differences in the number of authors of books in the top 50 genres. For example, whilst over 95% of romance novels are female-authored, males wrote 80% of the comic strips (Thelwall, 2017). There are less substantial differences in the average numbers of people that rate a book by author gender, although female-authored comic strips are particularly frequently rated. There are similar small gender differences in the average popularity of books by author gender, with female-authored comic strips and male authored literature receiving higher ratings (Thelwall, 2017).

Gender can also affect reviews of things other than books, showing that gender differences in cultural consumption and reception are widespread. People that watch TED Talks videos online can leave comments about them afterwards in YouTube or the TED website, and these can be short reviews. The comments tend to be more numerous and to express more sentiment if the presenter is female (Tsou et al., 2014), even though male-authored presentations tended to receive more Likes (Sugimoto et al., 2013). Perhaps some reviewers consciously or subconsciously support their own gender, empathize more readily with a same gender producer, are more interested in the topics covered, or enjoy the delivery style more readily.

Research questions

The objective of this study is to analyse the extent to which writer gender influences the attitudes of male and female readers to books within specific genres. It is important to look within genres for two reasons. First, since women and men read different types of books to some extent, an analysis that ignores genre risks inferring gender preferences from genre preferences. Second, genres are an important mechanism for differentiating between different types of books and so it is useful to obtain separate results for each one.

From the above reviewed work, there are gender differences in preferences for specific genres, but it is not known whether this occurs within genres. For example, do books in genres favoured by women attract more positive reviews from women than from the men that have also chosen to read a book from that genre? It is plausible that no differences would occur, given that both genders had (presumably) chosen a book within the genre to read. On the other hand, since females may be more positive in the social web (Thelwall et al., 2010) they may give higher ratings due to their preferred communicative style.

- RQ1: In which genres do female reviewers give higher ratings than male reviewers and vice versa?

Although reviewers tend to be more positive towards books authored by their own gender, does this occur once genre has been factored out and, if so, does it occur universally across genres? As discussed above, gender homophily does not seem to be a big factor in the social web and so this should not obscure book preference gender homophily.

- RQ2: In which genres are reviewer ratings biased towards books authored by the same gender?

One aspect of gender in book reviewing that does not seem to have been examined before is the relationship between reviewer gender and the reception of their review by other readers. For example, do readers empathize more with reviewers if they are from the same gender as the author or the reader, or does one gender tend to be more authoritative in the role of reviewer?

- RQ3: In which genres are female reviewers more liked than male reviewers and vice versa?
- RQ4: In which genres are reviews of books authored by the same gender as the reviewer more popular?

Finally, although there are theories about why different genders might enjoy specific types of books (e.g. Radway, 1984) there is little empirical evidence about how writer gender influences the way in which readers of different genders understand or engage with books in the genres that they have chosen to read.

- RQ5: Are there differences in the types of things that male and female reviewers write about male and female authored books in specific genres?

Methods

A large sample of amateur book reviews was needed to address the research questions. Goodreads was chosen as the source of the reviews because it appears to be the most popular book-based social website with public reviews. It was the 329th most visited website in the world in September 2016, according to Alexa.com (www.alexa.com/siteinfo/goodreads.com) and the 142nd most visited in the USA. Goodreads allows anyone to join and post reviews on any book, apparently without restrictions.

Books in Goodreads have a home page that contains brief information and lists reader reviews, if any. When there are more than 30 reviews about a book then only 30 are listed on the page, as selected by Goodreads with a secret algorithm designed to identify those that are most likely to interest readers. This algorithm presumably

exploits the popularity of each review, as judged by the number of its comments and/or likes from other members. Book reviews can therefore be sampled indirectly through their book rather than by author, reviewer or by date.

A random sample of 500,087 Goodreads books pages was selected as the raw material for the study. It is recycled from a previous article (Thelwall, 2017). The sample was chosen using a random number generator from an apparently exhaustive list of books from the Goodreads website map. A sample of 500,000 seemed to be enough to give substantial coverage of different genres. The final total is just over 500,000 because a larger starting list of 500,166 was used in case of deleted pages. The book pages were crawled using the free research web crawler SocSciBot (socscibot.wlv.ac.uk) and the reviews and metadata were extracted from the crawled pages using a function added to the free software Webometric Analyst (lexiurl.wlv.ac.uk). Each page contained up to 30 'community reviews'. For pages with more than 30 reviews in the website it is possible to change the selection of the most interesting to the most recent or oldest 30 but this option is not available to a web crawler because Goodreads uses a non-crawlable technology (an Ajax call). Thus, the reviews available for books are a poor quality sample because the selection method is an unknown Goodreads secret. This may have introduced a gender bias if, for example, the most interesting reviews were mainly written by females. Moreover, the selection applies disproportionately to female authors (51% of books with under 31 reviews had a female author and 62% with more). Nevertheless, only a minority of books (17%) had more than 30 reviews so any reviewer selection bias would be reduced by its relative scarcity within the dataset. For the main statistics, all reviews were kept for each book rather than just one per book because this is a natural way to bias the findings towards more reviewed books. It is a compromise between analysing on a per-book basis, for which one review should be selected per book, and a per-review basis, in which case all reviews should be analysed.

Only reviews of English-language books, as recorded on the Goodreads website, were included. Although most reviews seem to be in English, there are also non-English reviews for some English books. No attempt was made to filter these out for the text analysis, although many would have been filtered out by the gender identification stage described below. Automatic language identification is possible and reasonably accurate, but introduces errors and a bias because it is more accurate for longer texts. Hence its use would not be beneficial overall.

Author and reviewer genders were estimated using their first names, as recorded in Goodreads. First names were matched to genders using a list of 4772 common first names that are at least 90% male or at least 90% female in the USA 1990 census and had at least four letters. Authors and reviewers with first names not meeting this criterion

were discarded. This method is a practical technique to identify gender automatically but its small percentage of errors is likely to slightly reduce any gender differences found by swapping a small number of genders. The method assumes that both reviewers and authors usually record a name that reflects their gender, although there are some exceptions. These include names that have changed gender over time (e.g. Leslie), authors and reviewers that hide their gender (e.g. George Eliot), authors and reviewers that use a short version of their first name that changes their inferred gender (e.g. Ali Smith), and collectives that use a pseudonym (e.g. Nicci French, Nicolas Bourbaki). It also ignores genders that are not common enough to be analysed using the quantitative methods here. Each review was thus assigned to one of five categories: male author, female reviewer (MAFR); male author, male reviewer (MAMR); female author, male reviewer (FAMR); female author, female reviewer (FAFR); or other (author or reviewer gender unknown).

Books were assigned to a genre using the genre lists in the Goodreads book home pages, which contain the (up to) 10 most popular user-assigned book genres. A total of 201,560 books had at least one genre. Although book topics and structures are not fully constrained by genre conventions (Pyrhönen, 2007) these lists were used as a reasonable indication of book type. Each book was recorded as a member of the genres listed for it. Fiction lacks official genre classifications and so ad-hoc methods are needed to classify it (e.g. Harada, 2005) and reader judgements seem to be a reasonable choice. All (up to) 10 genres were included because some are very broad (e.g. Fiction) and some genres are subcategories of broader genres and so selecting only the top genre, which would also be a reasonable strategy, would tend to generate genres that are too broad. In any case, whichever strategy is chosen, it is impossible to remove the effect of genre blending (e.g. romance within fantasy) because genres are fluid and so no genre assignment strategy can perfectly separate them.

Genres are formally defined within Goodreads (Appendix, Table A2) but the proportion of members that are aware of these definitions and attempt to follow them is unknown. Members are allowed to allocate books to 'shelves' as an organizing device. Whilst some shelves are not genres, such as to-read and guilty pleasures, whenever a book is allocated to a shelf that is the name of a genre, such as romance or crime, this counts as one genre assignment for the book. Although there is a complete list of genres within the site (www.goodreads.com/genres/list) its origin is not explained. Since the genre names in the list are not standard (e.g. F-M-F) and exclude popular functional shelves (e.g. to-read) they appear to have been manually curated at least partly from popular shelf names created by Goodreads users. In some cases the shelf names are nested within the genre lists in the site. For example,

F-M-F is listed as Menage>F-M-F. This nesting is also not explained in the site and appears to be manually created by Goodreads staff rather than users or software.

All analyses were conducted separately for the 50 genres with the most books (see Appendix, Table A1 for numbers of books and ratings in each genre). It is important to analyse books within specific genres as far as possible because there are known gender differences in genre preferences that may influence comparisons between genders if multiple genres are included.

For the first four research questions, relevant statistics were extracted from the data. Confidence intervals are unreliable because the data is not independent – the same person could contribute multiple reviews, and many reviews would be included for individual popular books – but were calculated to give a very approximate indication of the variability in the data.

For the fifth research question, a word frequency comparison approach was used, with qualitative follow up investigations to seek evidence of gender differences in the content of texts for selected genres. For the word frequency comparison, each review in a given genre was processed to extract all words when both author and reviewer gender had been assigned. For each book a maximum of one review was selected at random. In addition, reviewers were selected a maximum of once. This ensures that the data is relatively independent since each review was from a different book and a different reviewer. This independence is necessary for the word frequency approach used because even a moderate amount of non-independence would cause problems with word frequency comparisons. Nevertheless, reviews were allowed for multiple books by the same writer, which allowed a degree of domination by individual prolific authors. This did not seem to be too large to justify adding an extra restriction for authors. In addition, the results were manually filtered for spam when checking the word frequency results. A common source of spam was a short author biography about a prolific author that had been reposted by different Goodreads users (sometimes librarians) to many books by that author.

Words occurring in male-authored reviews were compared to words occurring in female authored reviews and ranked using a difference between proportions test. For example, if the word 'hat' occurred in 15% of male authored reviews but 10% of female-authored reviews for a genre then the difference in proportions would be 5% and the difference in proportions test would assign a probability of this difference occurring by chance if there was no underlying gender difference. The top ranked disproportionately male terms were then examined qualitatively to assign an apparent cause and the same for the disproportionately female words. The process was repeated for male-authored books in comparison to female-authored books. This analysis was not carried out for all 50 genres because this would be unnecessary and

there was too little data in many cases, but only for a few genres in order to demonstrate convincingly that gender differences in review content *could* happen, rather than that they *always* happened.

Results

A list of the main genres extracted is given in Table 1 and genre descriptions from Goodreads are given in Appendix, Table A2.

RQ1: Reviewer gender differences in genres ratings

There are substantial differences in the average ratings given to books by male and female reviewers (Figures 1 and 2). Although males give higher average ratings for most genres (30 out of 50 have bars on the left in Figures 1 and 2), there are also many genres for which females give higher ratings. The two genres for which female reviewers gave the highest ratings relative to males are contemporary romance and paranormal romance. This fits the female preference for romances and the recent popularity of vampire romances. Despite this, males gave slightly higher ratings for romances overall, which is surprising. The two genres for which male reviewers gave the highest ratings relative to females are m-m-romance and short-stories.

Overall, then, it seems that male and female preferences for books cannot be completely explained by genre differences in preferences since they also occur within genres, including within relatively fine-grained genres such as paranormal romance.

RQ2: Reviewer-author gender homophily in book ratings

Reviewers rated authors of their own gender more highly in most genres (Figures 3 and 4). For female reviewers, the biggest same-gender differences are that they preferred manga and erotica written by other women. In contrast, male reviewers preferred romance and historical romance written by other men. Because the data includes all genres against which a book was classified, this may be due to genre blending. Men may tend to dislike romances, which tend to be written by women, and prefer other genres, such as mystery or adventure, which may blend an element of romance. This element of romance may lead mystery or adventure novels to be classified as romances by some readers, affecting the overall Goodreads classifications.

RQ3: Reviewer gender difference in review like counts

For almost all genres, male-authored reviews received more likes than did female-authored reviews (Figures 5

and 6). The text analysis below suggests that men tended to give more detailed reviews and this may be the explanation for males being apparently more authoritative, sympathetic or entertaining reviewers. Since the reason why people Like reviews in Goodreads is unknown, a definitive conclusion cannot be drawn from this difference.

RQ4: Reviewer-author gender homophily in review like counts

Goodreads users did not follow a strong pattern of giving more Likes to reviewers from the same gender as the author (Figures 7 and 8): members tended to Like reviews written by people of the opposite gender to the author. A possible reason is that these reviews might be more critical and hence longer – with long reviews being more likely to receive Likes. Alternatively, cross-gender reviews might be a minority within a genre and there would therefore be fewer to share the likes given to reviews expressing their opinion. Neither of these explanations is convincing, though.

RQ5: Word frequency case studies

This section analyses a sample of case studies designed to identify and confirm the existence of male and female differences in book reviewing. The broadest categories have been avoided (e.g. non-fiction) because these are the most likely to be affected by the subgenres chosen by readers and authors, and more gendered genres have been avoided because of the lack of reviews for minority genders (Table A1). Statistical details of the reviewer gender results are given in the Appendix, Table A1, to illustrate the calculations and only the ranked term lists are given below. To give some context, the most important concepts valued by fiction readers seem to be beauty and suspense (Knoop et al., 2016) but there do not seem to be gender differences in the use of terms related to these.

Mystery and fantasy. The mystery category had more female authors and reviewers than males but substantial numbers of both. The following points can be deduced from the results and additional examination of the data (Table 2).

- Romance in mysteries was important to both female authors and female reviewers (e.g. romance, love, loved).
- Relationships were important to both female authors and female reviewers (e.g. husband, mother, daughter, sister, friend – c.f. wife for male authors).
- Women were more important to female reviewers (e.g. Jane [referring to both authors and characters with this name], Margaret, woman).

Table 1. The number of books and ratings and identified male and female reviewers for the top 50 genres.

Genre*	Books	Ratings	Male Reviewers	Female reviewers	Reviews for RQ5
fiction	41,475	1,218,673	69470	220,826	5187
non-fiction	40,208	507,491	69,899	125,264	8215
romance	29,205	676,026	6805	119,519	3342
fantasy	19,909	1,057,426	26,409	68,596	2758
history	16,315	199,503	33,017	37,310	4033
children's	14,147	163,267	11,264	74,404	1989
contemporary	13,853	204,599	8471	72,730	227
mystery	13,093	389,375	20,210	72,440	3645
young-adult	11,286	621,919	10,739	60,915	1943
science-fiction	9967	261,253	22,221	22,363	1614
historical-fiction	9243	309,406	12,213	60,237	1909
fantasy>paranormal	9094	261,909	4556	34,374	599
historical	8654	137,803	12,514	57,776	260
sequential-art>comics	8567	166,331	13,334	9749	1263
biography	7925	103,156	18571	35,705	1627
short-stories	7834	96,615	8555	17,644	758
children's>picture-books	7410	131,850	4754	43,752	2945
romance>contemporary-romance	7403	91,478	868	30,965	212
adult	7043	72,240	7151	50,876	101
poetry	7011	111,621	5686	13,389	1943
adult-fiction>erotica	6981	78,255	906	13,487	427
sequential-art>graphic-novels	6961	169,828	13,204	10,828	878
sequential-art>manga	6623	285,353	349	1306	162
humor	6409	87,725	10,417	34,633	516
reference	6039	27,524	8862	16,453	580
romance>m-m-romance	5729	125,520	1100	5847	525
horror	5545	161,636	9923	14,398	914
classics	5187	664,000	10,818	18,831	556
philosophy	5131	95,606	11,234	7772	857
religion	5056	54,552	11,505	15,890	676
thriller	5003	86,473	12,521	26,326	453
adventure	4822	83,352	13,506	26,298	180
mystery>crime	4786	72,899	11,691	26,793	272
novels	4564	52,933	11,389	21,551	76
art	4469	30,879	4043	9718	876
science	4463	71,467	9908	14,006	938
christian	4356	45,478	7915	24,530	796
romance>paranormal-romance	4239	110,105	706	17,100	288
women's-fiction>chick-lit	4072	91,559	1318	35,144	481
politics	3894	34,030	12,657	9790	490
suspense	3829	41,560	6874	25,647	79
romance>historical-romance	3767	103,730	555	21,370	872
literature	3697	77,384	9679	13,979	92
autobiography>memoir	3673	67,055	8576	25,807	480
food-and-drink>cookbooks	3642	36,381	1183	7833	899
animals	3280	29,674	3501	24,264	294
psychology	3259	49,520	6378	12,558	617
thriller>mystery-thriller	3167	26,621	7562	25,407	30
fantasy>magic	3028	60,821	3188	15,762	70
travel	2941	31,811	4369	9781	654

*The symbol > indicates that the category on the right has been classified by Goodreads as being a subcategory of the category on the left.

The final column contains the number of independent reviews per genre for the top 50 genres – each book and each reviewer have a maximum of one review included in the set. Note that fiction and non-fiction are themselves genres under this scheme and so, for example, romance is not a subgenre of fiction here.

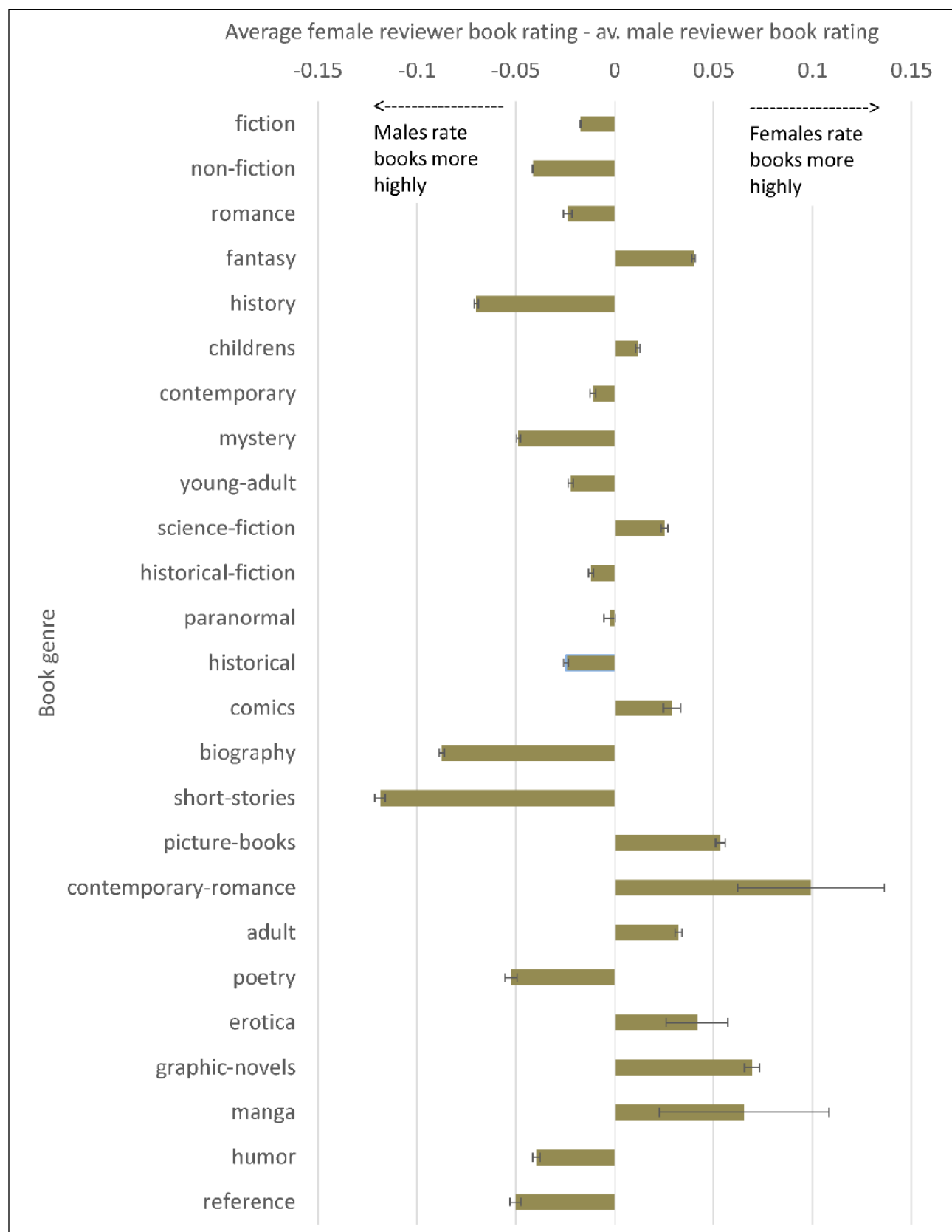


Figure 1. The average rating given to books by female reviewers subtract the average rating given to books by male reviewers for books in the top 25 genres.

Error bars give indicative approximate 95% confidence intervals.

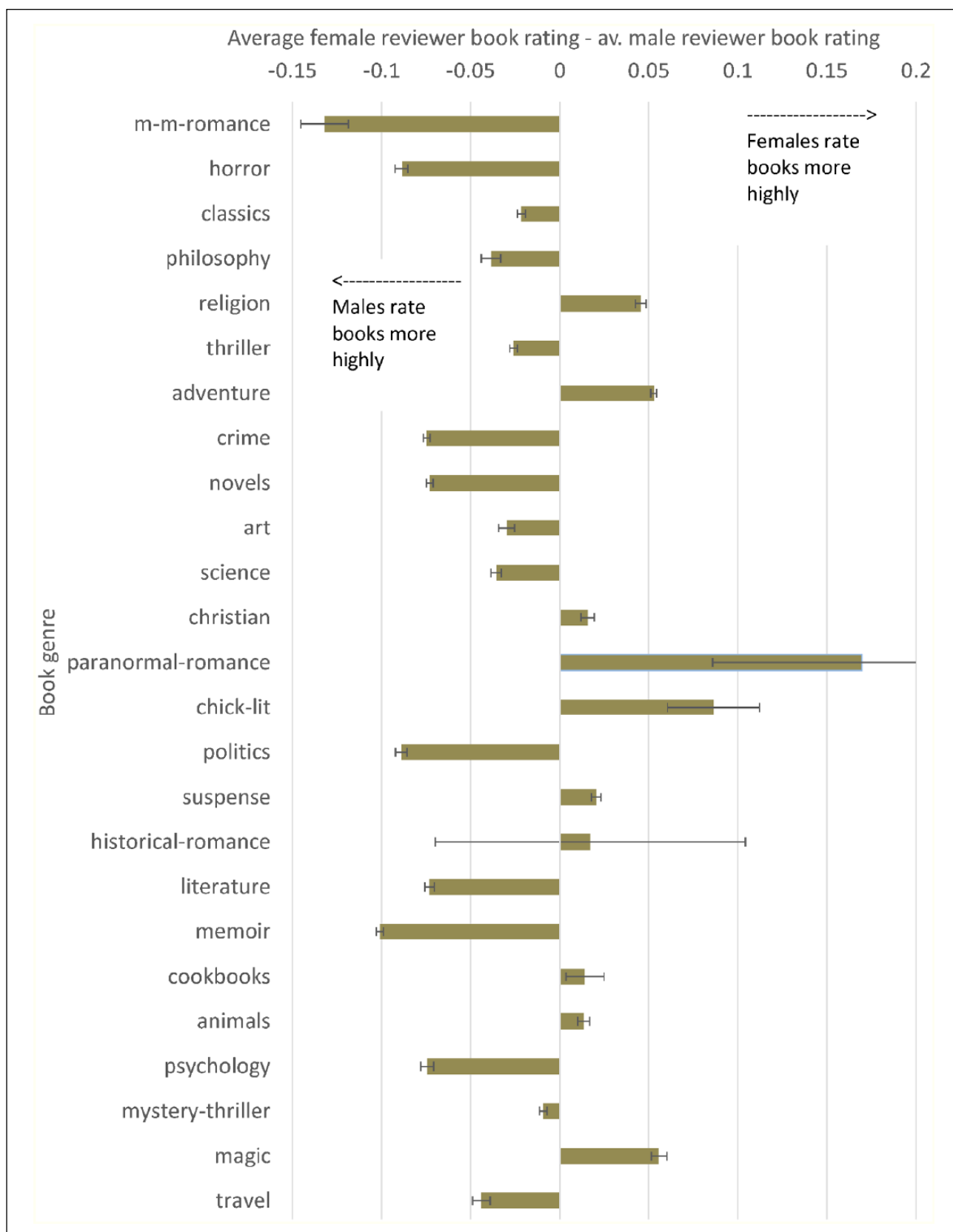


Figure 2. The average rating given to books by female reviewers subtract the average rating given to books by male reviewers for books in genres 26-50. Error bars give indicative approximate 95% confidence intervals.

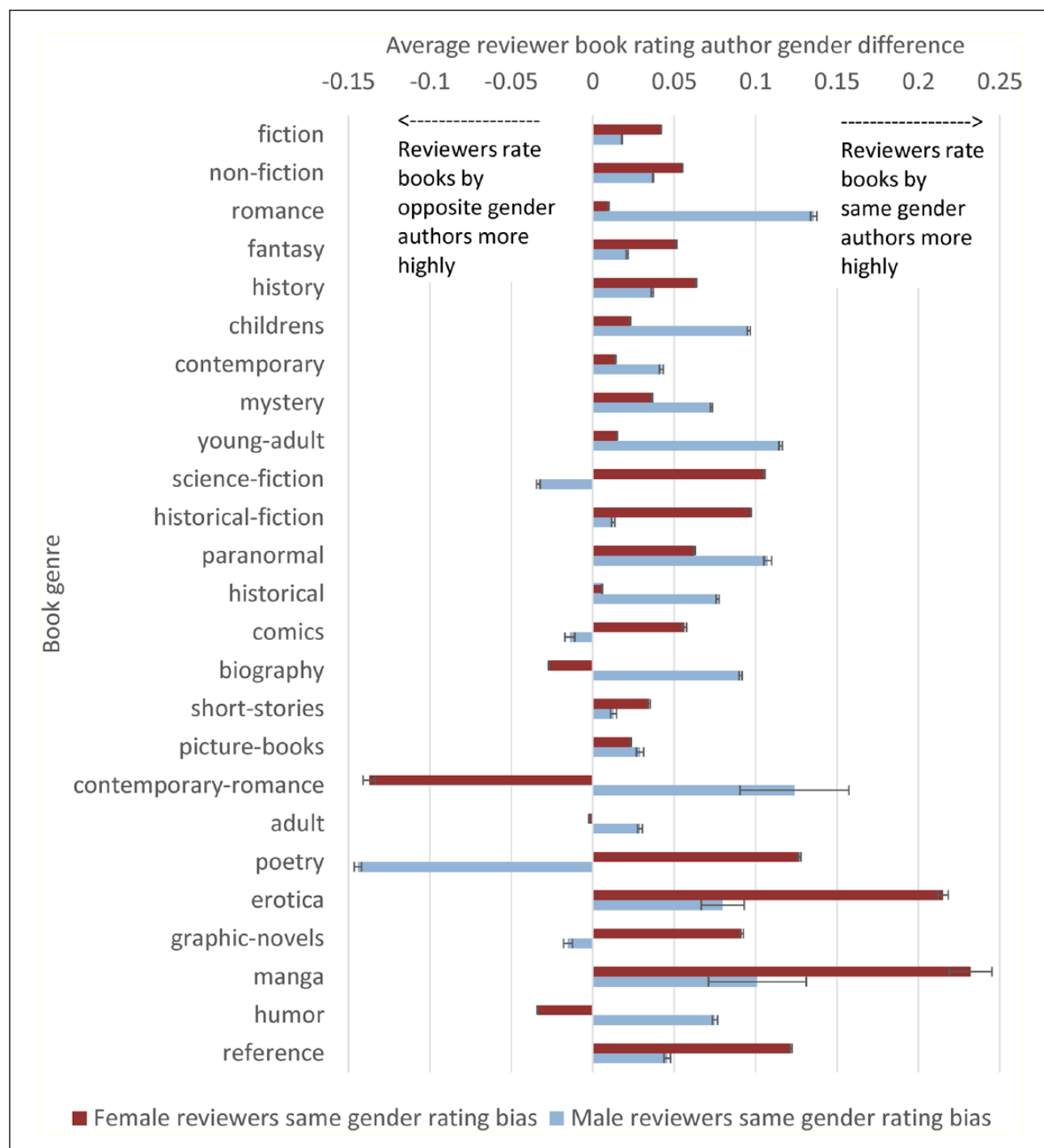


Figure 3. The average same gender rating bias (the average rating for books authored by the same gender, subtract the average rating for books authored by the opposite gender) for male and female reviewers of books in the top 25 genres. Error bars give indicative approximate 95% confidence intervals.

- Uncertainty was more important to female reviewers (e.g. mystery, guessing, guess).
- Structure and style were more important to male reviewers (e.g. narrative, prose, noir, extended, pre-tentious – c.f. cozy [the cozy crime genre] for female authors).

- Male authors' books were more likely to have been converted into screenplays (e.g. film, Hollywood).

The results for fantasy echo those for mystery, and especially the importance of romance for female reviewers and authors (Table 3).

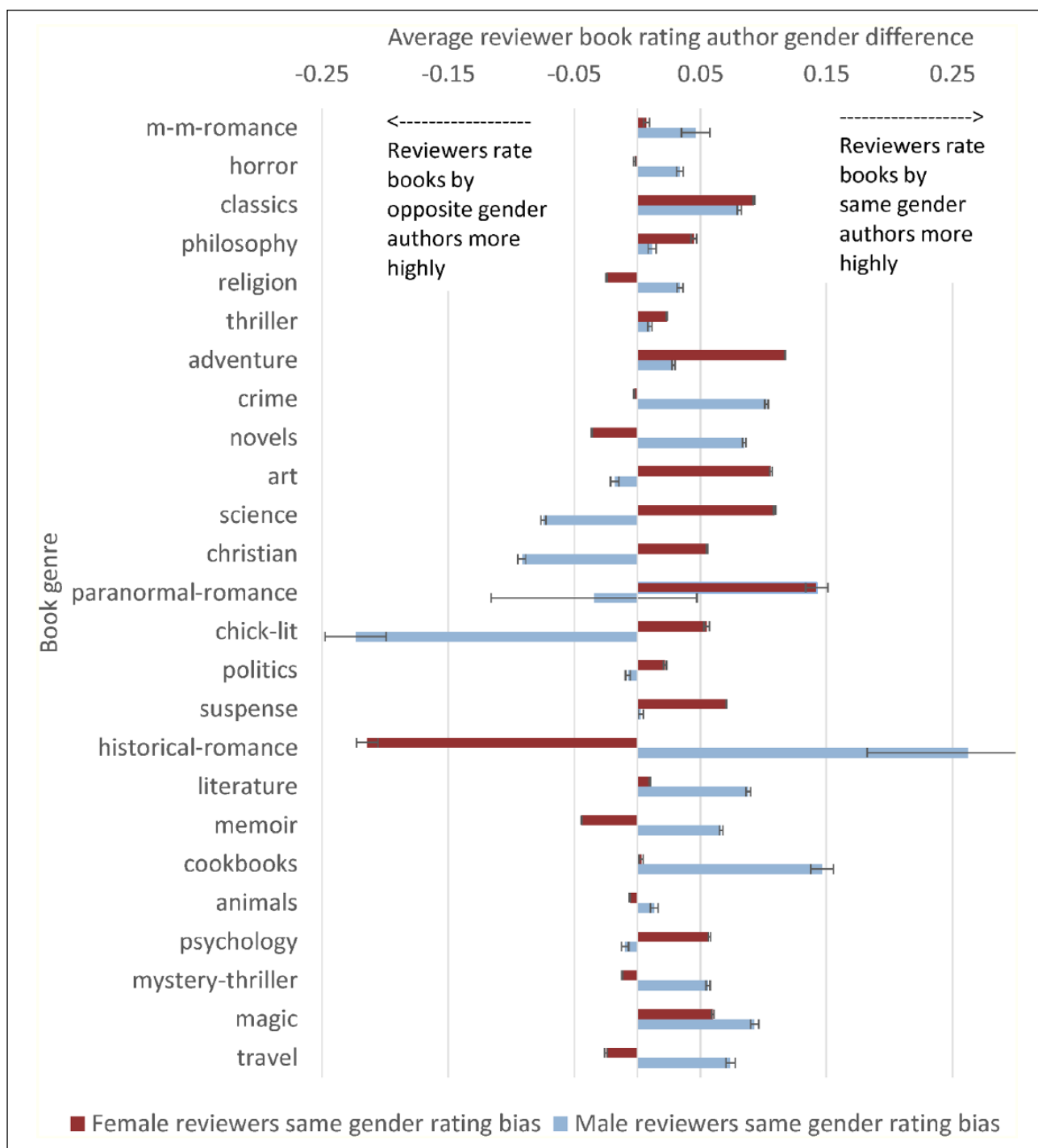


Figure 4. The average same gender rating bias (the average rating for books authored by the same gender, subtract the average rating for books authored by the opposite gender) for male and female reviewers of books in genres 26–50. Error bars give indicative approximate 95% confidence intervals.

Comics. A little can be deduced about gender differences in the comics category but there were so few female authors that the results are highly speculative (Table 4). Female reviewers and authors tended to focus on positive stories (e.g. love, gift, beautifully, and perhaps also girly; although pretty is a male author term) whereas male reviewers

focused more on superheroes and crime (e.g. superman, punisher).

M-m-romance. The m-m-romance category results are speculative due to the small numbers involved (Table 5). Some repeated text was removed before processing,

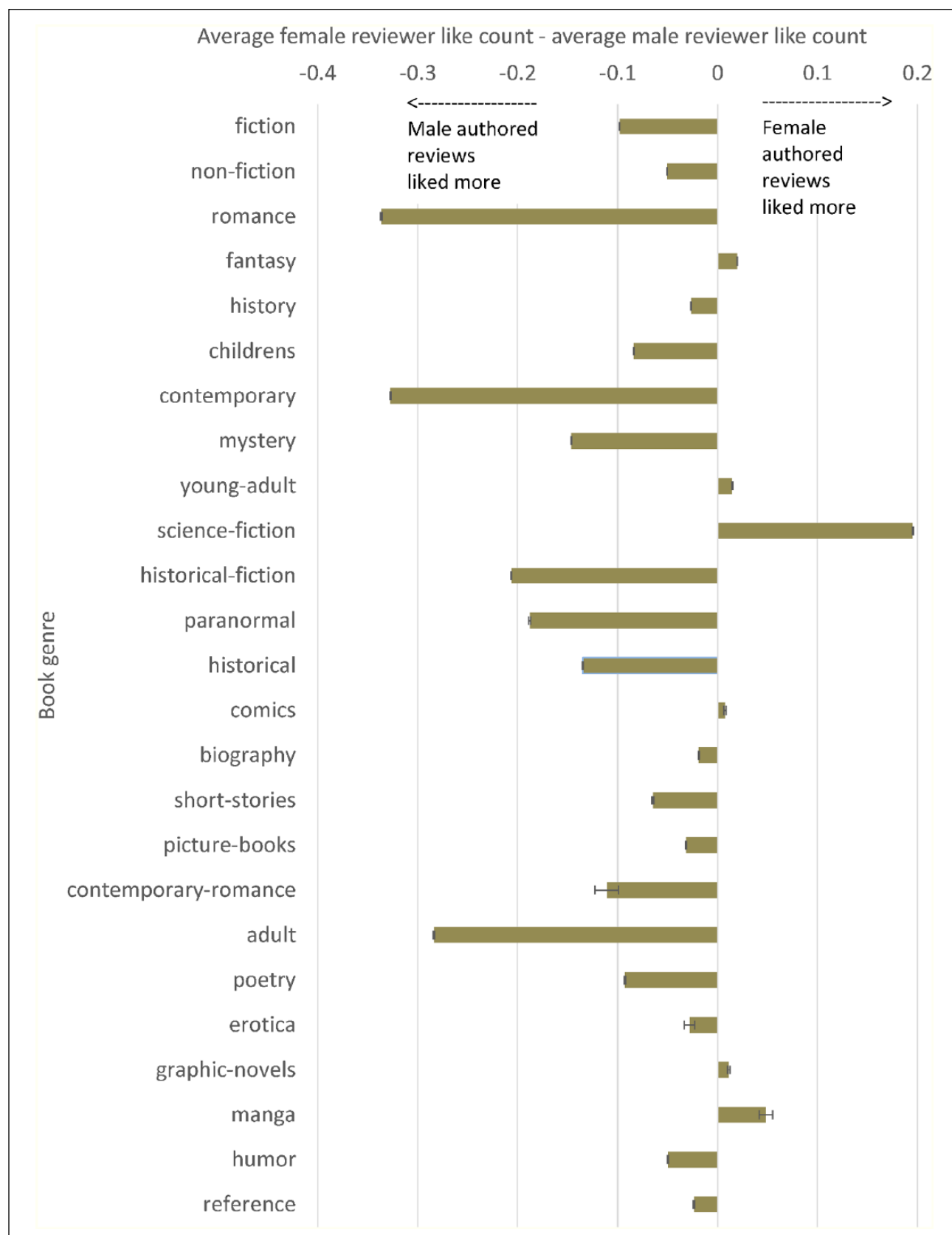


Figure 5. The average number of likes for reviews written by females subtract the (geometric mean) average number of likes for reviews written by males for the top 25 genres. Error bars give indicative approximate 95% confidence intervals.

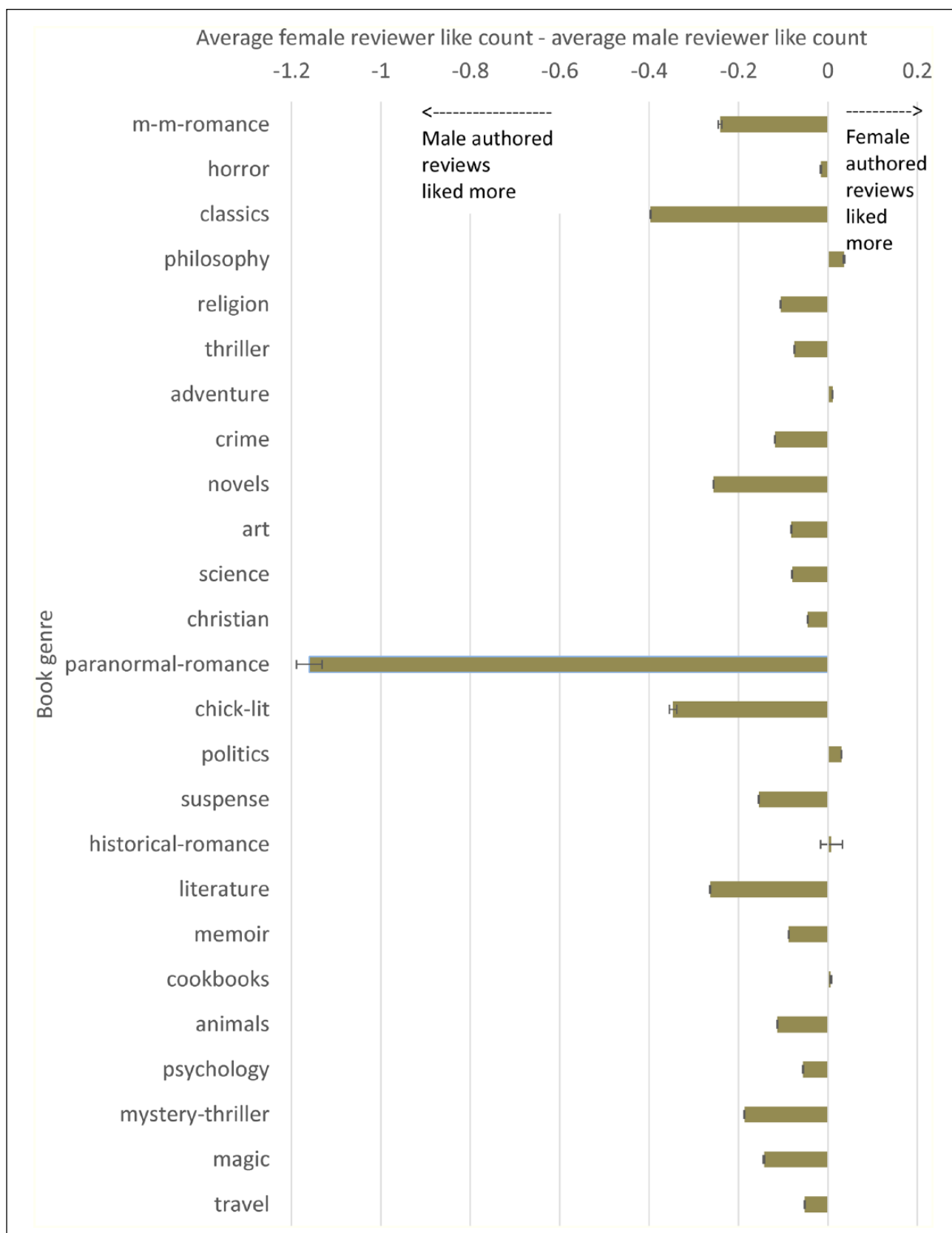


Figure 6. The average number of likes for reviews written by females subtract the (geometric mean) average number of likes for reviews written by males for genres 26–50.

Error bars give indicative approximate 95% confidence intervals.

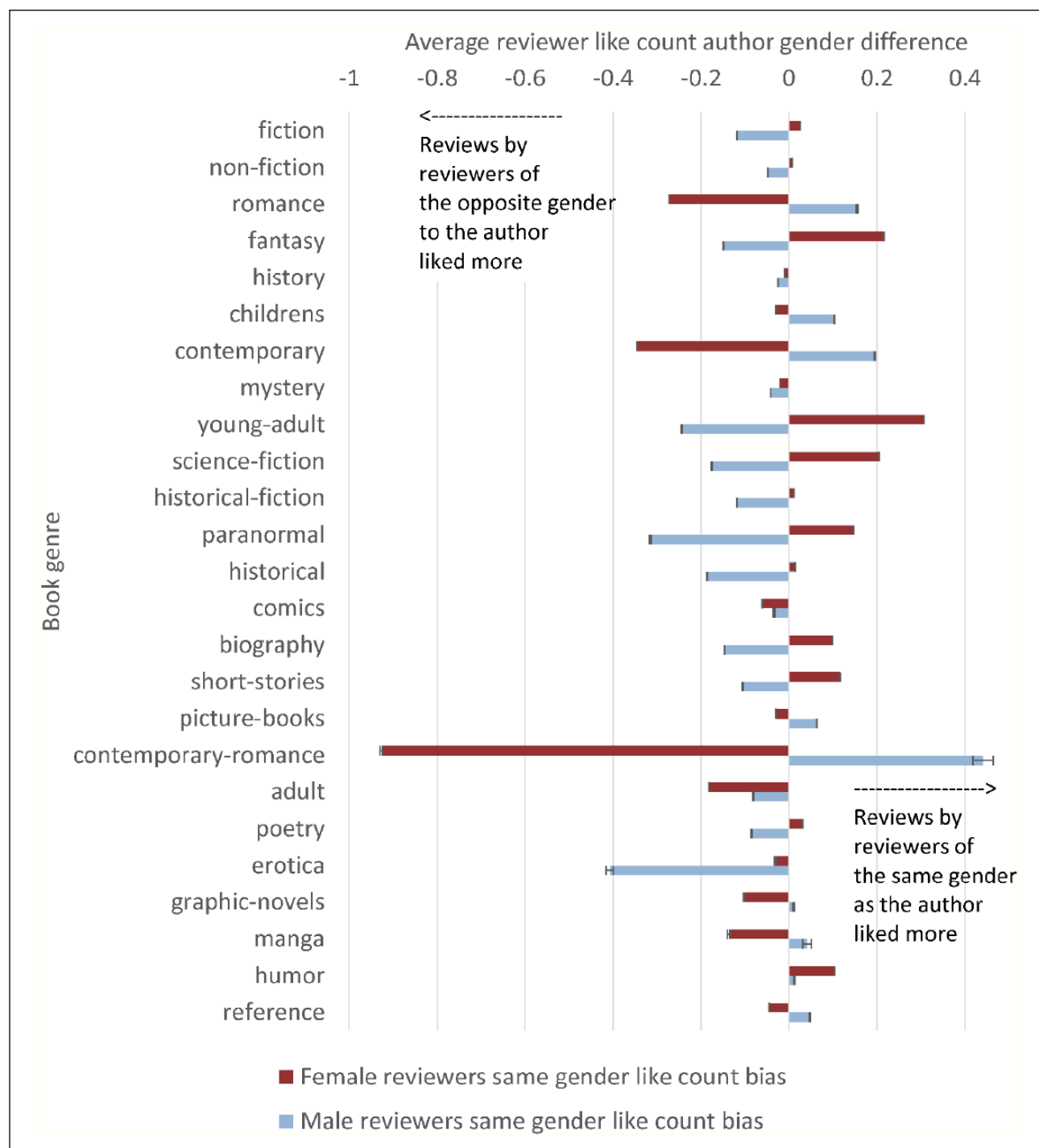


Figure 7. The average same gender like bias (the average number of likes for books reviewed by the same gender, subtract the average rating for books authored by the opposite gender) for male and female reviewers of books in the top 25 genres. Error bars give indicative approximate 95% confidence intervals.

‘Librarian Note: There is more than one author in the Goodreads database with this name. See this thread for more information. erotic & gay romances’.

- Female reviewers tended to be positive about aspects of the books (e.g. the terms: nice, liked, recommend).
- Female authors tended to include more fantasy elements (e.g. shifter, mate, angel).
- Female authors tended to write sequences of books (e.g. series).
- Male authors tended to be discussed in more political terms (e.g. gay, homosexual).

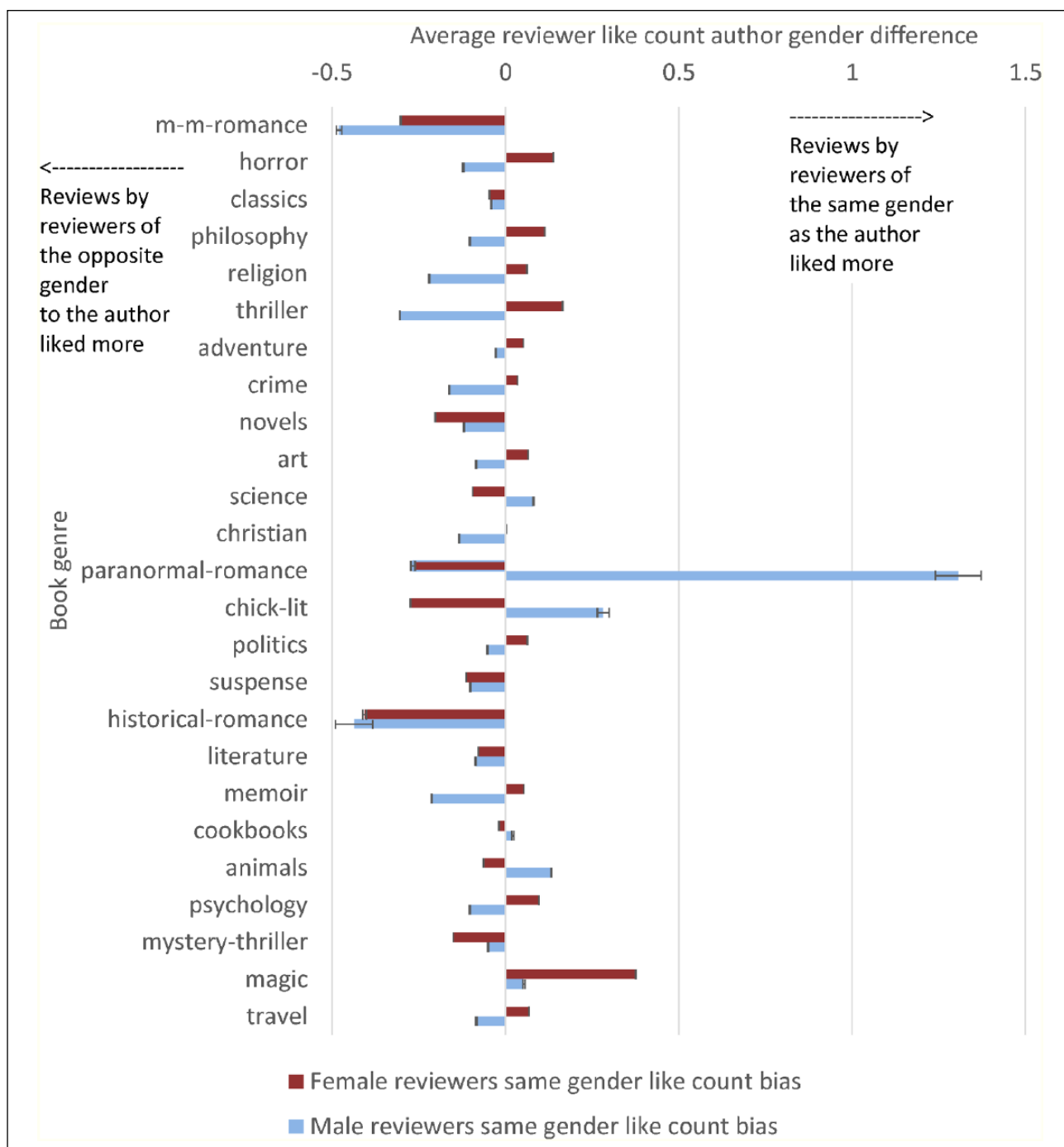


Figure 8. The average same gender like bias (the average number of likes for books reviewed by the same gender, subtract the average rating for books authored by the opposite gender) for male and female reviewers of books in genres 26–50. Error bars give indicative approximate 95% confidence intervals.

- Male reviewers focused more on a detailed review of the book structure (crime, character, predictable).

Limitations

An important limitation is that cultural consumption and preferences vary by social class and occupation as well

as age (Bihagen and Katz-Gerro, 2000) and these factors have not been taken into account in the analysis. Some of the differences found could therefore be due to different age ranges or social classes reading genres from the different genders. The reviewers and raters are self-selected and so it is not clear that their opinions reflect those of typical book readers. Moreover, members

Table 2. The top 20 terms in terms of difference between proportions z value for reviewers and authors by gender for mystery. 3493 reviews in category, MAFR:972; MAMR:489; FAMR:200; FAFR:1832.

Rank	Female reviewer	Male reviewer	Female author	Male author
1	she	novel	she	him
2	love	Chandler	husband	novel
3	loved	Raymond	herself	John
4	husband	certainly	cozy	David
5	guessing	scudder	mother	Sherlock
6	romance	Hardy	romance	Holmes
7	woman	perhaps	Nancy	James
8	Jane	John	mystery	Robert
9	liked	hero	sister	film
10	mystery	prose	heroine	Michael
11	French	narrative	Ms	Patterson
12	Grace	noir	cat	database
13	exchange	Spenser	Jane	Chandler
14	daughter	extended	lady	Raymond
15	wasn't	here	love	Spenser
16	seat	exception	loved	wife
17	national	roughly	Sarah	Hollywood
18	recipe	pretentious	friend	Hardy
19	guess	Charlie	Drew	royal
20	Margaret	BA	Kate	Dick

Table 3. The top 20 terms in terms of difference between proportions z value for reviewers and authors by gender for fantasy. 2662 reviews in category, MAFR:745; MAMR:654; FAMR:191; FAFR:1072.

Rank	Female reviewer	Male reviewer	Female author	Male author
1	she	Conan	she	john
2	love	Robert	romance	Mr
3	loved	Howard	love	Robert
4	children	novel	loved	comic
5	fairy	D&D	husband	Conan
6	want	American	girl	himself
7	know	que	herself	best
8	beautiful	game	romantic	David
9	mother	pulp	exchange	de
10	romance	David	magic	wife
11	boy	Moorcock	Jane	war
12	eye	volume	cat	literary
13	happen	recent	heroine	Howard
14	gave	Leiber	can	least
15	exchange	story-telling	mother	Charles
16	because	imprint	paranormal	Stephen
17	friend	original	Diana	Michael
18	heart	epic	oh	pulp
19	friendship	Fritz	connection	Paul
20	magic	historia	heart	pseudonym

exploit Goodreads and other book-based websites for different purposes (Thelwall and Kousha, 2017; Vlieghe et al., 2016) and so the gender differences in ratings and reviews may partly reflect gender differences in styles of using the site.

The results may be affected by small numbers of commercial reviews and reviews ghost written by book authors. The presence of reviews written in English by non-native speakers may influence the word frequency results in a different way. Presumably simpler terms in

Table 4. The top 20 terms in terms of difference between proportions z value for reviewers and authors by gender for the comics genre.

1234 reviews in category, MAFR:364; MAMR:806; FAMR:26; FAFR:38. The total differs from Table 2 due to omitting null reviews.

Rank	Female reviewer	Male reviewer	Female author	Male author
1	female	stories	Simone	star
2	she	Kirby	referred	after
3	laugh	seem	Gail	well
4	recently	nature	Lynda	both
5	deadpool	superman	class	around
6	love	five	husband	action
7	wanting	sense	illuminate	include
8	wasn't	these	Louise	pretty
9	I'm	Jack	Pook's	man
10	Alice	punisher	Comeek	through
11	gift	will	girly	reader
12	women	complex	refrigerator	keep
13	haven't	event	Sfar	took
14	suit	artist	weezie	far
15	understand	mark	zenescope	became
16	beautifully	work	you'll	wrote
17	didn't	after	well-known	year
18	referenced	solid	Barry	him
19	cute	Stan	poorly	creator
20	staff	most	bird	Eisner

Table 5. The top 20 terms in terms of difference between proportions z value for reviewers and authors by gender for m-m-romance.

494 reviews in category, MAFR:106; MAMR:47; FAMR:35; FAFR:306.

Rank	Female reviewer	Male reviewer	Female author	Male author
1	nice	crime	she	Mr
2	liked	available	shifter	gay
3	recommend	race	mate	entertaining
4	ended	further	contemporary	late
5	care	skimmed	husband	young
6	few	manner	Ms	novel
7	parent	predictable	gorgeous	whose
8	highly	carefully	series	give
9	taking	final	publishing	trip
10	off	determined	bought	partner
11	brother	collection	plus	complaint
12	opinion	fun	disappointed	homosexual
13	home	street	despite	shadow
14	rushed	character	cat	grey
15	exactly	winning	angel	mystery
16	couldn't	obviously	girl	any
17	sequel	behavior	stories	began
18	meet	rape	necessary	date
19	note	sale	generally	finished
20	rating	serve	took	followed

categories that are most adopted by non-native speakers would have their frequencies boosted by this. Most significantly, the results are likely to be substantially

affected by male and female differences in book genres overlapping with genre differences. Thus, gender differences for a genre may be due to gendered preferences for

overlapping genres. Since books rarely fit neatly into genres, this is an unavoidable problem.

The issue of preferences is also somewhat problematic. In this paper, a person registering a book in the site for a genre has been taken as evidence of a preference for that genre. This may not always be true because books can be registered and reviewed for other purposes, such as class assignments, recommendations from friends or alignment with TV series, rather than selected with an expectation of enjoyment. It changes the interpretation of the results if some of the books reviewed were not selected by the reader but were on their course or school syllabus. In any case, social pressure might also exert a strong persuasion on individuals, such as teenage girls feeling that they really ought to read the Twilight novels (Aubrey et al., 2010 – and some leaving very negative reviews as a result).

The dominance of female users of Goodreads is a potential source of bias because it seems likely that a higher proportion of female book readers use Goodreads. Although most of the results split males from females and so are not directly affected by this disparity, they will be indirectly affected by having a wider female reader sample. The findings for male and female reviewers may therefore contrast female readers with a sample of male readers that are relatively more enthusiastic about books. This is purely speculative, however, given that nothing is known about why more female than male readers join Goodreads.

As mentioned above, another limitation for the analyses using sampled reviews is that only the 30 most interesting were available for the 17% of books that had at least 31 reviews. The selection process for these reviews may have introduced biases.

The results apply mainly to a US context due to the popularity of Goodreads within this country. There are substantial differences in author genders and the types of books that men and women tend to write in other nations, such as a smaller number of female writers in France (Verboord, 2012). Other countries, such as China (Qian et al., 2008) have even more different traditions or the roles of female authors. Thus, the findings do not necessarily apply outside of the USA.

A statistical limitation is that whilst differences have been found and appear to be statistically significant, as judged by the approximate confidence intervals on the graphs, these confidence intervals are only guidelines due to the non-independence of the data and are also very narrow due to the huge sample sizes involved. Moreover, it is not clear how important the differences found are. All rating differences are substantially less than a whole rating category and so are relatively small. For example, the results fall far short of showing that all male romance readers 'didn't like' (one star) them but all female readers thought they were 'amazing' (five stars). Judgements like this are complicated by the tendency of most reviewers to

give relatively high ratings. It is also not clear whether the differences found are due to differing subjective judgements about the books or whether the differences are caused by other factors, such as attempts to conform to gender roles within a public space. Perhaps males enjoying romances might avoid reviewing them or pretend to dislike them in their reviews. Similarly, male trolls might post scathing reviews of popular female genre books, such as paranormal romances, with the goal of generating a reaction.

Finally, the analysis here has not attempted to deconstruct how people take pleasure from reading (e.g. Zunshine, 2006). This might give more insights into why gender differences in preferences occur and how they may be subtly shaped by wider social forces.

Because of the many limitations of this study, the findings must be interpreted cautiously. Nevertheless, they have value in giving a new type of insight into the relationship between readers, writers and genre. This is especially important given the lack of systematic evidence of this important relationship from other sources.

Discussion

In answer to the first research question, there are differences in the average ratings of books within individual genres according to whether the reader was male or female (at least those who wrote a review). In contrast to a social network study, women were not more positive overall (Thelwall et al., 2010). Male and female differences in genre preferences amongst book readers were previously known about (e.g. women preferring romances) but the existence of reader differences in preferences for books *within* genres of those who presumably chose to read them is a new finding. It is perhaps unsurprising that women rate books that are classified in one of the three heterosexual romance sub-genres (contemporary-romance, historical-romance and/or paranormal-romance) more highly than do men, given that women are the overwhelming majority of romance readers. This outcome is not a foregone conclusion since the reviewers are presumably the people that chose to read the book and therefore the male romance readers would be exceptions for their gender. Thus, it seems that even men that like romance sub-genre novels enough to read them like them, on average, less than do their female counterparts. The opposite is true for the romance genre as a whole, however. An explanation for this might be that the romance genre is often assigned to books as an additional genre classification because they have a romantic element, whereas romance sub-genres tend to include books that fit well within the sub-genre (e.g. paranormal romance). Alternatively, readers that have a single main genre preference may classify their books by their secondary genre (as a main genre) to avoid having all of their books on a single shelf. In contrast,

male reviewers seem to give higher ratings to more serious genres, including non-fiction, history, biography, poetry and reference. Taking the example of romances, perhaps males and females are randomly influenced by numerous gender-specific social influences that lead to female readers on average liking romances more than male readers. The males that read romances would then be outliers for their gender, perhaps only just working up the enthusiasm to read a romance, whereas similarly extreme females would absolutely love them. If true, this would lead to the average for males being lower than the average for female readers.

For the second research question, reviewer ratings are biased towards books authored by the same gender within nearly all genres. Whilst reviewer bias towards authors of the same gender has been previously noticed in the case of male reviewers favouring male-authored works, this has been explained in terms of female genres being ignored. The results extend this knowledge by showing that reviewer same gender preferences exist within genres, extend beyond professional reviewers to amateur reviewers and are not universal across all genres. The gender preference within genres could occur because the reader likes the style of authors of their own gender (see the above mentions of: Brooks, 1992; Winnett, 1990), irrespective of the topic. Alternatively, authors of the same gender may show more detailed knowledge of, or display more interest in, topics of common interest with the reader (e.g. sports, female fashions) or topics within the reader's worldview (e.g. breastfeeding, testicular cancer). Authors of the same gender may also include aspects of other genres favoured by their own gender within their works, even if they write within other genres (e.g. romantic themes in some Harry Potter books). Whatever the causes, this finding suggests that the historical gender bias of the literary elite discussed above extends more widely to the book-reading public – not just in terms of the books selected to read but also for their opinions of books read.

Male-authored reviews tended to be more liked than female-authored reviews in most genres, with science-fiction being the main exception. This may be due to male reviewers leaving more detailed 'serious' reviews overall (e.g. using terms like *narrative*, *prose* and *noir* in Table 2, *pulp* and *imprint* in Table 3 and *sense* and *complex* in Table 4, with no similar female terms occurring in Tables 2–5). There may be an element of schadenfreude in some of the categories since the Likes did not necessarily originate from other book readers. For example, the apparent extreme popularity of male reviews of paranormal romance may be due to a few leaving bad reviews of books that annoyed them because significant females in their lives (e.g. girlfriend, sister) vocally praised them. One of the most popular reviews of one of the most popular books in this genre was a sarcastic attack by a male who had shelved it in 'books-that-i-fucking-hate'.

Although there were similar female-authored reviews of this book, there was a higher proportion of very enthusiastic reviews from female readers.

In answer to the fourth research question, reviews of books authored by the same gender as the reviewer were not systematically more popular. There was a slight tendency for the opposite to be true, but the reason for this is unclear.

For the fifth research question, the comparison of the frequencies of words in male-authored reviews with those of female-authored reviews for individual genres found plausible evidence of gender-specific themes, such as romance in mysteries and fantasy for female authors and reviewers, and structure and style for male reviewers. This is presumably due to people singling out plot aspects that appeal to them or that are important to them in reviews, even if that is not the main book genre. Whilst the results are unsurprising, they confirm the general importance of gender-targeting themes, such as romance and relationships, within non-gender-specific genres, such as mystery and fantasy.

Conclusions and practical implications

The results show that gender differences run deep for book readers, at least considering those who write reviews in Goodreads. These differences are not restricted to the preferences of types of books to read but also occur within genres. Thus, for example, female amateur book reviewers rate contemporary romance more highly than do male amateur reviewers, even though both groups have decided to read this genre. In addition, within almost all genres, readers prefer books written by the same gender author. This within genre gender preference may be due to authors being more likely to mention people of their own gender (male writers foregrounding women has been historically rare) or incorporate gender-specific knowledge or topics (e.g. relationships, war) that may bore the opposite gender. Thus, in the future writers and publishers targeting both genders may wish to consider strategies to ensure that neither males nor females are alienated from their work by a lack of interest in the details. Similarly, librarians recommending books may also take these factors into account.

The almost universal greater popularity of male-authored reviews may be an artefact of females using the site more for informal communication (Thelwall and Kousha, 2017), and hence may not help to understand the relationship between reading and gender for books.

The results also show that Goodreads data can be used to extract gender information about the authors and reviewers of books. Presumably reviewer opinions broadly reflect reader opinions to some extent, although online reviewers seem likely to be a gender and education biased subset of readers. The numerical data has practical value for

publishers, librarians and booksellers to give them insights into the market segments that books target. The word frequency results can also give information about the themes appreciated by the different genders within the same genre. Of course, these should not be interpreted as evidence for a formula for authors and publishers to follow because highly successful books often break genre conventions or create new genres. Even for an author wishing to break genre, however, the results might give ideas about how a particular genre currently works and hence the conventions that are available to be broken.

At the methodological level, the word frequency approach followed here was successful but points to the need for careful filtering of reviews to remove spam and reduce the possibility of undue influence for individual books and reviewers. A consequence of this is that a huge amount of data is needed for a reasonable analysis. In the current paper, despite starting with over half a million books, there was insufficient data for a word frequency analysis for many of the most popular genres. Future research may attempt to extract larger samples in order to obtain more fine-grained results. Nevertheless, large scale review text analysis seems to provide a useful complement to the growing area of macroanalyses of the contents of books (e.g. Jockers, 2013). Together, they may give new insights into how genres evolve, how the public reacts to the different aspects of genres, and gender differences in the reception of books. In addition, the results for each individual genre would be interesting to follow up in their own right for a more fine-grained analysis, as would the finding that reviews by readers of the opposite gender to the author tend to be more liked. Finally, Goodreads itself is a valuable resource about the reception of books by readers. A recent call for the creation of a corpus of this kind of information (Boot, 2013) has therefore been answered as a by-product of a successful commercial site.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Aubrey JS, Walus S and Click MA (2010) Twilight and the production of the 21st century teen idol. In: Click MA (ed.) *Bitten by Twilight: Youth Culture, Media, & the Vampire Franchise*. New York: Peter Lang, pp. 225–242.
- Bamman D, Eisenstein J and Schnoebelen T (2014) Gender identity and lexical variation in social media. *Journal of Sociolinguistics* 18(2): 135–160.
- Baym N (1987) *Novels, Readers, and Reviewers: Responses to Fiction in Antebellum America*. New York: Cornell University Press.
- Berkers P, Janssen S and Verboord M (2011) Globalization and ethnic diversity in western newspaper coverage of literary authors: Comparing developments in France, Germany, the Netherlands, and the United States, 1955 to 2005. *American Behavioral Scientist* 55(5): 624–641.
- Beutel ME, Brähler E, Glaesmer H, et al. (2011) Regular and problematic leisure-time Internet use in the community: Results from a German population-based survey. *Cyberpsychology, Behavior, and Social Networking* 14(5): 291–296.
- Bihagen E and Katz-Gerro T (2000) Culture consumption in Sweden: The stability of gender differences. *Poetics* 27(5): 327–349.
- Brewer J (2013) *The pleasures of the Imagination: English Culture in the Eighteenth Century*. London: Routledge.
- Boot P (2013) The desirability of a corpus of online book responses. In: *Proceedings of the workshop on computational linguistics for literature*, Atlanta, GA, USA, 9–14 June 2013, pp. 32–40). Stroudsburg, PA: Association for Computational Linguistics (ACL).
- Boyd D (2007) Why youth (heart) social network sites: The role of networked publics in teenage social life. In: Buckingham D (ed.) *MacArthur Foundation Series on Digital Learning—Youth, Identity, and Digital Media Volume*. Cambridge, MA: MIT Press, pp. 119–142. Available at: <https://www.danah.org/papers/WhyYouthHeart.pdf> (accessed 8 May 2017).
- Brooks P (1992) *Reading for the Plot: Design and Intention in Narrative*. Cambridge, MA: Harvard University Press.
- Bureau of Labor Statistics (2015) Time spent in leisure activities in 2014, by gender, age, and educational attainment. *The Economics Daily*. Available at: <http://www.bls.gov/opub/ted/2015/time-spent-in-leisure-activities-in-2014-by-gender-age-and-educational-attainment.htm> (accessed 10 September 2016).
- Casey EM (1996) Edging women out?: Reviews of women novelists in the 'Athenaeum' 1860–1900. *Victorian Studies* 39(2): 151–171.
- Castle T (1997) Women and literary criticism. In: Nisbet H and Rawson C (eds) *The Cambridge History of Literary Criticism IV: The Eighteenth Century*. Cambridge: Cambridge University Press, pp. 434–455.
- Charles EB (2011) *Literary Criticism: An Introduction to Theory and Practice*. 5th edn. London: Longman.
- Cherney ID and London K (2006) Gender-linked differences in the toys, television shows, computer games, and outdoor activities of 5-to 13-year-old children. *Sex Roles* 54(9–10): 717–726.
- Cherny L (2012) UK bestsellers: Remash by genre and gender. In: *Ghostweather R&D Blog*. Available at: <http://blogger.ghostweather.com/2012/08/uk-bestsellers-remash-by-genre-and.html> (accessed 8 May 2017).
- Chevalier JA and Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research* 43(3): 345–354.
- Chong PK (2013) Legitimate judgment in art, the scientific world reversed? *Social Studies of Science* 43(2): 265–281.
- Chong PK (2015) Playing nice, being mean, and the space in between: Book critics and the difficulties of writing bad reviews. In: Berthoin A, Hutter M and Stark D (eds) *Moments of Valuation: Exploring Sites of Dissonance*. Oxford: OUP, pp. 133–146.

- Cleland ND (1984) Comparison of sports coverage in *Book Review Digest* and *Book Review Index*. *RQ* 23(4): 451–459.
- Cushman G, Veal AJ and Zuzanek J (eds) (2005) *Free Time and Leisure Participation: International Perspectives*. Wallingford: CABI.
- d'Astous A, Colbert F and Mbarek I (2006) Factors influencing readers' interest in new book releases: An experimental study. *Poetics* 34(2): 134–147.
- Ekelund BG and Börjesson M (2002) The shape of the literary career: An analysis of publishing trajectories. *Poetics* 30(5): 341–364.
- Ellison NB, Steinfield C and Lampe C (2007) The benefits of Facebook 'friends': Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication* 12(4): 1143–1168.
- Ferriss S and Young M (eds) (2013) *Chick Lit: The New Woman's Fiction*. New York: Routledge.
- Foster CL (1951) Some thoughts concerning book reviewing. *Educational Research Bulletin* 30(3): 64–66.
- Gallagher C (1988) Who was that masked woman? The prostitute and the playwright in the comedies of Aphra Behn. *Women's Studies: An Interdisciplinary Journal* 15(1/3): 23–42.
- Goldin C, Katz LF and Kuziemko I (2006) The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives* 20(4): 133–133.
- Griswold W, McDonnell T and Wright N (2005) Reading and the reading class in the twenty-first century. *Annual Review of Sociology* 34(1): 127–141.
- Guillory J (2013) *Cultural Capital: The Problem of Literary Canon Formation*. Chicago, IL: University of Chicago Press.
- Harada T (2005) Toward automatic classification of novels using kansei keywords in book reviews. *Japan Society of Information and Knowledge* 15(2): 57–62. [原田隆史 (2005). 書評中の感性キーワードを用いた小説の分類 (第13回 (2005年度) 研究報告会講演論文集). 情報知識学会誌 15(2): 57–62.]
- Harsh C (2001) Reviewing new woman fiction in the daily press: *The Times*, the *Scotsman*, and the *Daily Telegraph*. *Victorian Periodicals Review* 34(1): 79–96.
- Hartmann T and Klimmt C (2006) Gender and computer games: Exploring females' dislikes. *Journal of Computer-Mediated Communication* 11(4): 910–931.
- Henning B and Vorderer P (2001) Psychological escapism: Predicting the amount of television viewing by need for cognition. *Journal of Communication* 51(1): 100–120.
- Hipsky M (2011) *Modernism and the Women's Popular Romance in Britain, 1885–1925*. Columbus, OH: Ohio University Press.
- Hooper B (2010) *Writing Reviews for Readers' Advisory*. Chicago, IL: American Library Association.
- Horning KT (2010) *From Cover to Cover: Evaluating and Reviewing Children's Books*. Rev edn. New York: Harper Collins.
- Hutner H (ed.) (1993) *Rereading Aphra Behn: History, Theory, and Criticism*. Charlottesville, VA: University of Virginia Press.
- Isaacson N (2013) Science fiction for the nation: Tales of the moon colony and the birth of modern Chinese fiction. *Science Fiction Studies* 40(1): 33–54.
- Jackson P (2008) Can a man really write a Mills & Boon? *BBC News*. Available at: <http://news.bbc.co.uk/1/hi/7710508.stm> (accessed 8 May 2017).
- Janssen S (1997) Reviewing as social practice: Institutional constraints on critics' attention for contemporary fiction. *Poetics* 24(5): 275–297.
- Janssen S (1998) Side-roads to success: The effect of sideline activities on the status of writers. *Poetics* 25(5): 265–280.
- Jockers ML (2013) *Macroanalysis: Digital Methods and Literary History*. Urbana, IL: University of Illinois Press.
- Knoop CA, Wagner V, Jacobsen T, et al. (2016) Mapping the aesthetic space of literature 'from below'. *Poetics* 56(1): 35–49. DOI: 10.1016/j.poetic.2016.02.001
- Kuei Huang Y and Yang WI (2010) A study of internet book reviews and borrowing intention. *Library Review* 59(7): 512–521.
- Lauer JD (1989) Urgency and overlap in the dissemination of information: Five religion journals and the book reviews they published in 1986. *Collection Management* 11(3–4): 113–126.
- Lin TM, Luarn P and Huang YK (2005) Effect of Internet book reviews on purchase intention: A focus group study. *Journal of Academic Librarianship* 31(5): 461–468.
- Little J (2014) *On the Importance of Pseudonymous Activity*. Available at: <http://dearauthor.com/features/essays/on-the-importance-of-pseudonymous-activity/> (accessed 8 May 2017).
- Lizardo O (2006) The puzzle of women's 'highbrow' culture consumption: Integrating gender and work into Bourdieu's class theory of taste. *Poetics* 34(1): 1–23.
- McGinty S and Moore AC (2008) Role of gender in reviewers' appraisals of quality in political science books: A content analysis. *Journal of Academic Librarianship* 34(4): 288–294.
- McPherson M, Smith-Lovin L and Cook JM (2001) Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–444.
- Milota M (2014) From 'compelling and mystical' to 'makes you want to commit suicide': Quantifying the spectrum of online reader responses. *Scientific Study of Literature* 4(2): 178–195.
- Monthly Review (1844) *The Monthly Review from January to May Inclusive*. London: G. Henderson.
- Moore M (1978) Discrimination or favoritism? Sex bias in book reviews. *American Psychologist* 33(10): 936–938.
- Moreno AI and Suárez L (2008) A study of critical attitude across English and Spanish academic book reviews. *Journal of English for Academic Purposes* 7(1): 15–26.
- Mudambi SM and Schuff D (2010) What makes a helpful review? A study of customer reviews on Amazon. com. *MIS Quarterly* 34(1): 185–200.
- Muscannell NL and Guadagno RE (2012) Make new friends or keep the old: Gender and personality differences in social networking use. *Computers in Human Behavior* 28(1): 107–112.
- Nelson DD (2001) Women in public. In: Bauer D and Gould P (eds) *The Cambridge Companion to Nineteenth-Century American Women's Writing*. Cambridge: Cambridge University Press, pp. 38–68.
- Ortega y Miranda E (1996) On book reviewing. *Journal of Educational Thought* 30(2): 191–202.
- Ortiz MH (2015) *New York Times* bestsellers: A look at the authors. In: *Martin Hill Ortiz Author's Blog*. Available at: <http://martinhillortiz.blogspot.co.uk/2015/05/new-york-times-bestsellers-look-at.html> (accessed 8 May 2017).

- Osmani R (2015) Writing a book review in the EFL classroom: Ways to enhance reading comprehension and writing skills. In: Anon *Writing and Critical Approaches to Literature*. Macedonia: SEEU-Tetovo, pp. 159–166.
- Palmer JW (1987) An inquiry into the availability of Canadian fiction in US libraries with special attention to the influence of reviews. *Library Acquisitions: Practice & Theory* 11(4): 283–295.
- Palmer JW (1991) Fiction selection in Ontario public libraries: How important are reviews? *Public Library Quarterly* 10(4): 39–48.
- Palmer JW (1995) Agreement and disagreement among fiction reviews in *Library Journal*, *Booklist*, and *Publishers Weekly*. *Library Resources & Technical Services* 39(4): 384–390.
- Posner RA (1992) When is parody fair use? *Journal of Legal Studies* 21(1): 67–78.
- Prescott S (2003) *Women, Authorship and Literary Culture, 1690–1740*. Basingstoke: Palgrave Macmillan.
- Pyrhönen H (2007) Genre. In: Herman D (ed.) *The Cambridge Companion to Narrative*. Cambridge: Cambridge University Press, pp. 109–123.
- Qian N, Fong GS and Smith RJ (eds) (2008) *Different Worlds of Discourse: Transformations of Gender and Genre in late Qing and early Republican China*. Leiden, The Netherlands: Koninklijke Brill NV.
- Radway JA (1984) *Reading the Romance: Women, Patriarchy, and Popular Literature*. Chapel Hill, NC: University of North Carolina Press.
- Rashley LH (2007) Revisioning gender: Inventing women in Ursula K. Le Guin's nonfiction. *Biography* 30(1): 22–47.
- Rees EL (2003) *Margaret Cavendish: Gender, Genre, Exile*. Manchester: Manchester University Press.
- Roberts JJ (2015) Amazon sues people who charge \$5 for fake reviews. *Fortune*. Available at: <http://fortune.com/2015/10/19/amazon-fake-reviews/> (accessed 8 May 2017).
- Rothbauer PM and McKechnie LE (2000) The treatment of gay and lesbian fiction for young adults in selected prominent reviewing media. *Collection Building* 19(1): 5–16.
- Smith A (2009) *The Internet as a Diversion*. Pew Research Center Available at: <http://www.pewinternet.org/2009/09/10/the-internet-as-a-diversion/> (accessed 8 May 2017).
- Söderberg J (1998) Consumption, gender and preferences in Sweden, 1920–1965. *Scandinavian Economic History Review* 46(1): 71–84.
- Stanton JP (1988) Statistical profile of women writing in English from 1660 to 1800. In: Keener FM and Lorsch SE (1988) *Eighteenth-century Women and the Arts*. New York: Greenwood, pp. 247–254.
- Steiner A (2008) Private criticism in the public space: Personal writing on literature in readers' reviews on Amazon. *Participations* 5(2). Available at: http://www.participations.org/Volume%205/Issue%202/5_02_steiner.htm (accessed 8 May 2017).
- Stutzman F, Capra R and Thompson J (2011) Factors mediating disclosure in social network sites. *Computers in Human Behavior* 27(1): 590–598.
- Sugimoto CR, Thelwall M, Larivière V, et al. (2013) Scientists popularizing science: Characteristics and impact of TED talk presenters. *PloS one* 8(4): e62403.
- Tapper O (2014) Romance and innovation in twenty-first century publishing. *Publishing Research Quarterly* 30(2): 249–259.
- Tepper SJ (2000) Fiction reading in America: Explaining the gender gap. *Poetics* 27(4): 255–275.
- Thelwall M (2008) Social networks, gender and friending: An analysis of MySpace member profiles. *Journal of the American Society for Information Science and Technology* 59(8): 1321–1330.
- Thelwall M (2009) Homophily in MySpace. *Journal of the American Society for Information Science and Technology* 60(2): 219–231.
- Thelwall M (2017) Book genre and author gender in Goodreads: Romance>paranormal-romance to autobiography>memoir. *Journal of the Association for Information Science and Technology* 68(5): 1212–1223. DOI:10.1002/asi.23768
- Thelwall M and Kousha K (2017) Goodreads: A social network site for book readers. *Journal of the Association for Information Science and Technology* 68(4): 972–983. DOI: 10.1002/asi.23733
- Thelwall M, Wilkinson D and Uppal S (2010) Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the American Society for Information Science and Technology* 61(1): 190–199.
- Todorov T (1990) *Genres in Discourse*. Cambridge: Cambridge University Press.
- Tse P and Hyland K (2009) Discipline and gender: Constructing rhetorical identity in book reviews. In: Hyland K and Diani G (eds) *Academic Evaluation: Review Genres in University Settings*. London: Palgrave-MacMillan, pp. 105–121.
- Tsou A, Thelwall M, Mongeon P, et al. (2014) A community of curious souls: An analysis of commenting behavior on TED talks videos. *PloS one* 9(4): e93609.
- Tuchman G and Fortin NE (2012) *Edging Women Out: Victorian Novelists, Publishers and Social Change*. London: Routledge.
- Tufekci Z (2008) Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society* 28(1): 20–36.
- Turner C (1992) *Living by the Pen: Women Writers in the Eighteenth Century*. London: Routledge.
- Ugander B, Karrer B, Backstrom L, et al. (2011) The anatomy of the facebook social graph. *arXiv arXiv:1111.4503*.
- Upton L (1991) Mourning monsters: Deception and transformation in Rachel Ingalls's fiction. *Critique: Studies in Contemporary Fiction* 33(1): 53–61.
- Van Rees CJ (1987) How reviewers reach consensus on the value of literary works. *Poetics* 16(3): 275–294.
- Van Rees CJ (1989) The institutional foundation of a critic's connoisseurship. *Poetics* 18(1): 179–198.
- Verboord M (2003) Classification of authors by literary prestige. *Poetics* 31(3): 259–281.
- Verboord M (2009) The legitimacy of book critics in the age of the Internet and omnivorosity: Expert critics, Internet critics and peer critics in Flanders and the Netherlands. *European Sociological Review* jcp039. DOI:10.1093/esr/jcp039
- Verboord M (2011) Cultural products go online: Comparing the Internet and print media on distributions of gender, genre and commercial success. *Communications* 36(4): 441–462.
- Verboord M (2012) Female bestsellers: A cross-national study of gender inequality and the popular–highbrow culture divide

- in fiction book production, 1960–2009. *European Journal of Communication* 27(4): 395–409.
- VIDA (2015) *The 2014 VIDA Count*. Available at: <http://www.vidaweb.org/2014-vida-count/> (accessed 8 May 2017).
- Vlieghe J, Muls J and Rutten K (2016) Everybody reads: Reader engagement with literature in social media environments. *Poetics* 54(1): 25–37.
- Wallace DP (1987) Some thoughts on book reviewing. *Journal of the American Society for Information Science* 38(4): 227–227.
- Wheeler Carlo P, Duchin D and Natowitz A (1998) Reviewing American history: Coverage and timeliness in *American Historical Review*, *Choice*, and *Journal of American History*. *Collection Building* 17(2): 71–79.
- Winnett S (1990) Coming unstrung: Women, men, narrative, and principles of pleasure. *Publications of the Modern Language Association of America* 103(3): 505–518.
- Xu Y and Xi Y (2013) Ethical literary criticism: Its openness, inclusiveness and vitality. *Foreign Literature Studies* 35(6): 171–176.
- Young EV (1993) Aphra Behn, gender, and pastoral. *Studies in English literature, 1500–1900* 33(3): 523–543.
- Zunshine L (2006) *Why We Read Fiction: Theory of Mind and the Novel*. Columbus, OH: Ohio State University Press.

Author biography

Mike Thelwall is professor of Information Science and leader of the Statistical Cybermetrics Research Group at the University of Wolverhampton, UK, which he joined in 1989. He is also Docent at the Department of Information Studies at Åbo Akademi University and a research associate at the Oxford Internet Institute. He investigates web phenomena using quantitative-led research methods and provides a range of free software for research.

Appendix

Table A1. Complete details of the word frequency comparison between female and male reviewers for the Mystery genre. Top 20 words for female reviewers (MAFR+FAFR) compared to male reviewers (MAMR+FAMR).

Term	Overall term freq.	Female-authored reviews using term	Female-authored reviews using term (%)	Male-authored reviews using term	Male-authored reviews using term (%)	Difference in proportions z value
she	9750	23%	86	12%	11%	6.2
love	8911	15%	55	8%	7%	4.8
loved	4166	6%	17	2%	4%	4
husband	1763	4%	13	2%	3%	3.1
guessing	237	3%	5	1%	2%	3
romance	2646	3%	9	1%	2%	2.8
woman	2101	5%	18	3%	2%	2.6
Jane	479	2%	2	0%	1%	2.6
liked	3319	7%	33	5%	3%	2.4
mystery	2161	29%	170	25%	4%	2.3
French	653	1%	1	0%	1%	2.3
grace	333	1%	0	0%	1%	2.3
exchange	447	1%	0	0%	1%	2.3
daughter	1533	3%	10	1%	1%	2.2
wasn't	2070	5%	20	3%	2%	2.2
seat	210	1%	1	0%	1%	2.2
national	934	1%	1	0%	1%	2.2
recipe	786	1%	0	0%	1%	2.2
guess	922	3%	8	1%	1%	2.2
Margaret	259	1%	1	0%	1%	2.2

Table A2. Genre definitions copied and pasted from Goodreads.

Genre* and definition

fiction Fiction is the telling of stories which are not real. More specifically, fiction is an imaginative form of narrative, one of the four basic rhetorical modes.

non-fiction Non-fiction is an account or representation of a subject which is presented as fact.

romance According to the Romance Writers of America, 'Two basic elements comprise every romance novel: a central love story and an emotionally-satisfying and optimistic ending'. Both the conflict and the climax of the novel should be directly related to that core theme of developing a romantic relationship, although the novel can also contain subplots that do not specifically relate to the main characters' romantic love. Other definitions of a romance novel may be broader, including other plots and endings or more than two people, or narrower, restricting the types of romances or conflicts.

fantasy Fantasy is a genre that uses magic and other supernatural forms as a primary element of plot, theme, and/or setting. Fantasy is generally distinguished from science fiction and horror by the expectation that it steers clear of technological and macabre themes, respectively, though there is a great deal of overlap between the three (collectively known as speculative fiction or science fiction/fantasy).

history History (from Greek ἱστορία - historia, meaning 'inquiry, knowledge acquired by investigation') is the discovery, collection, organization, and presentation of information about past events. History can also mean the period of time after writing was invented.

childrens Children's literature is for readers and listeners up to about age 12. It is often illustrated. The term is used in senses that sometimes exclude young-adult fiction, comic books, or other genres. Books specifically for children existed at least several hundred years ago.

contemporary Contemporary literature is literature with its setting generally after World War II.

mystery Mystery fiction is a loosely-defined term that is often used as a synonym of detective fiction – in other words a novel or short story in which a detective (either professional or amateur) solves a crime. The term 'mystery fiction' may sometimes be limited to the subset of detective stories in which the emphasis is on the puzzle element and its logical solution (cf. whodunit), as a contrast to hardboiled detective stories which focus on action and gritty realism. However, in more general usage 'mystery' may be used to describe any form of crime fiction, even if there is no mystery to be solved.

Table A2. (Continued)

young-adult Young-adult fiction [] is fiction written for, published for, or marketed to adolescents and young adults, roughly ages 13 to 18.

science-fiction Science fiction (abbreviated SF or sci-fi with varying punctuation and capitalization) is a broad genre of fiction that often involves speculations based on current or future science or technology.

historical-fiction Historical fiction presents a story set in the past, often during a significant time period. In historical fiction, the time period is an important part of the setting and often of the story itself.

fantasy>paranormal Paranormal books involve unusual experiences that lack a scientific explanation. Some popular subjects in paranormal books are supernatural creatures, ESP, clairvoyance, ghosts, UFOs, telepathy and psychics.

historical Works of fiction or non-fiction about the past.

sequential-art>comics A comic book or comicbook, also called comic magazine or simply comic, is a publication that consists of comic art in the form of sequential juxtaposed panels that represent individual scenes. Panels are often accompanied by brief descriptive prose and written narrative, usually dialog contained in word balloons emblematic of the comics art form.

biography A biography [] is an account of a person's life, usually published in the form of a book or essay, or in some other form, such as a film.

short-stories A short story is a short work of prose fiction. It may be in any genre of fiction, and the exact definition of 'short' will often depend on the genre.

childrens>picture-books A picture book combines visual and verbal narratives in a book format, most often aimed at young children.

romance>contemporary-romance Romance with a contemporary setting. Contemporary romance is set after World War II.

adult For adults. May sometimes contain themes considered unsuitable for children or teens.

poetry Poetry is a form of literary art in which language is used for its aesthetic and evocative qualities in addition to, or in lieu of, its apparent meaning.

adult-fiction>erotica Erotica is any story that places sex at the center of the story.

sequential-art>graphic-novels A graphic novel is a narrative work in which the story is conveyed to the reader using sequential art in either an experimental design or in a traditional comics format. The term is employed in a broad manner, encompassing non-fiction works and thematically linked short stories as well as fictional stories across a number of genres.

sequential-art>manga Japanese or Japanese-influenced comics and graphic novels. Usually printed in black-and-white. There are many genres inside manga, the most distinct being shojo (for girls) and shonen (for boys).

humor A comic novel is usually a work of fiction in which the writer seeks to amuse the reader, sometimes with subtlety and as part of a carefully woven narrative, sometimes above all other considerations.

reference A reference, is derived from Middle English referren, from Middle French référer, from Latin referre, 'to carry back' formed from a prefix re- and ferre, 'to bear'.

romance>m-m-romance M/M Romance is fiction with an intimate relationship between two (or more) men at its center. It could be contemporary, fantasy, sci-fi etc.

horror Horror fiction is fiction in any medium intended to scare, unsettle, or horrify the audience.

classics A classic stands the test of time. The work is usually considered to be a representation of the period in which it was written; and the work merits lasting recognition. In other words, if the book was published in the recent past, the work is not a classic.

philosophy Philosophy is the study of general problems concerning matters such as existence, knowledge, truth, beauty, justice, validity, mind and language.

religion Religion is a cultural system that establishes symbols that relate humanity to spirituality and moral values.

thriller Thrillers are characterized by fast pacing, frequent action, and resourceful heroes who must thwart the plans of more-powerful and better-equipped villains. Literary devices such as suspense, red herrings and cliffhangers are used extensively.

adventure Adventure fiction is a genre of fiction in which an adventure, an exciting undertaking involving risk and physical danger, forms the main storyline.

mystery>crime The crime genre includes the broad selection of books on criminals, courts, and investigations. Mystery novels are usually placed into this category.

novels A novel is a long prose narrative that usually describes fictional characters and events in the form of a sequential story.

art Books that showcase particular types of art.

science Science [] is the effort to discover, and increase human understanding of how the physical world works.

(Continued)

Table A2. (Continued)

christian The term 'Christian' is used adjectivally to describe anything associated with Christianity, or in a proverbial sense 'all that is noble, and good, and Christ-like'.

romance>paranormal-romance Paranormal romance is a sub-genre of the romance novel related to paranormal fiction. Paranormal romance focuses on romance and includes elements beyond the range of scientific explanation, blending together themes from the genres of urban fantasy, traditional fantasy, science fiction, or horror.

women's-fiction>chick-lit Chick lit is genre fiction which addresses issues of modern womanhood, often humorously and lightheartedly. Although it sometimes includes romantic elements, chick lit is generally not considered a direct subcategory of the romance novel genre, because the heroine's relationship with her family or friends is often just as important as her romantic relationships.

politics Politics [], is a process by which groups of people make collective decisions.

suspense Suspense is a feeling of uncertainty and anxiety about the outcome of certain actions, most often referring to an audience's perceptions in a dramatic work.

romance>historical-romance Historical romance is any genre romance set before 1945. They can take place in any location or period of history.

literature Literature is the art of written works. The two most basic written literary categories include fiction and non-fiction, although 'literature' in popular use can also mean a sub-genre of fiction called literary fiction.

autobiography>memoir [] Memoir is autobiographical writing, but not all autobiographical writing follows the criteria for memoir. Memoirs are structured differently from formal autobiographies which tend to encompass the writer's entire life span, focusing on the development of his/her personality.

food-and-drink> cookbooks Non-fiction books that contain a collection of recipes, techniques, and tricks of the trade or else focus on the exploration of food, cooking, and culture of food.

animals Animals are multicellular, eukaryotic organisms of the kingdom Animalia [].

psychology Books that involve psychology; the study of mental processes and human behavior.

thriller>mystery-thriller Mystery-thrillers are stories with twists and turns that keep you on your toes guessing about the truth, the POV's truth, and about what will happen next. They can be written from a hero's POV or from a sinister character's POV.

fantasy>magic These are all books that deal with magic: wands, spells, curses, potions, hexes, thoughts (mind-magic), art, symbols, rituals, magical languages, witchcraft, etc.

travel Travel is the movement of people or objects (such as airplanes, boats, trains and other conveyances) between relatively distant geographical locations. [] Travel writing is a genre that has, as its focus, accounts of real or imaginary places.

All were obtained by appending the genre name to the URL root <https://www.goodreads.com/genres/> and extracting the first part of the definition (ignoring any subsequent discussion). For example, the first genre is defined at URL <https://www.goodreads.com/genres/fiction>. Some definitions are taken from Wikipedia, such as that for animal, and some define a topic rather than a genre (e.g. animal, philosophy).

*The symbol > indicates that the category on the right has been classified by Goodreads as being a subcategory of the category on the left.