A Statistical Analysis of Editorial Influence and Author–Character Similarities in 1990s *New Yorker* Fiction

Katherine L. Milkman
Harvard University, Boston, MA 02138, USA

René Carmona and William Gleason
Princeton University, Princeton, NJ 08544, USA

Abstract

We present a quantitative analysis of 442 pieces of fiction published between 5 October 1992 and 17 September 2001 in the *New Yorker* magazine. We address two independent questions using the same data set. First, we examine whether changes in the Executive Editor or Fiction Editor are associated with significant changes in the type of fiction published at the *New Yorker*. Second, we examine whether *New Yorker* authors write fiction more often than not about characters with whom they share demographic traits. We find that changes in Fiction Editor at the *New Yorker* are associated with numerous significant, quantifiable changes in the magazine's fiction and that these effects are greater than those associated with a change in the *New Yorker*'s Executive Editor. We also find that authors of *New Yorker* fiction write significantly more often than not about protagonists who share their race, gender, and country of origin and who are within or below their age range. The same is true of secondary characters except in the case of gender.

Correspondence:
Katherine L. Milkman
E-mail:
kmilkman@fas.harvard.edu

1 Introduction

This article employs the tools of statistics to analyze various aspects of the fiction published in the *New Yorker* magazine between 5 October 1992 and 17 September 2001. The analyses conducted seek to answer two very different questions about *New Yorker* fiction from this era, both of which can be addressed using the same data set. The first question this article attempts to answer is whether changes in Executive Editors or Fiction Editors are associated with significant changes in the type of fiction published at the *New Yorker*. The second question is whether *New Yorker* authors write fiction about characters who are similar to themselves.

The *New Yorker* is the most prestigious, widely read and frequently published literary magazine in the world. Its influence on America's cultural scene is generally agreed to be unparalleled among publications of its genre, and since the early twentieth century, aspiring as well as established writers have viewed inclusion in its pages as one of the highest honors attainable, offering the potential to launch a literary career. In the past decade, the magazine has published an average of 52 pieces of fiction annually, selected through a process that necessarily involves subjective editorial judgment but is intended to be meritocratic. These stories are viewed by literary critics and writers worldwide as a sample of the most compelling modern short

fiction, although it should be noted that the *New Yorker* does not necessarily publish a representative sample of modern fiction. However, in light of its unique characteristics as a literary magazine that has consistently published what most would argue to be the best contemporary short fiction in a periodical, the *New Yorker* is an ideal venue for analyzing modern short stories.

Of particular interest to the authors of this article is the nature of the relationship between the *New Yorker*'s editors and the magazine's fictional content. Since the submissions editors inspire and the stories they subsequently choose to print significantly shape the universe, at least domestically, of published short fiction in any given month, year, or decade, gaining insight into what type of influence, if any, editors exert over the *New Yorker*'s fiction content seems a worthwhile endeavor. For example, if the personal preferences of editors—conscious or unconscious—or pressures to satisfy particular audiences appear to influence what fiction is printed in the *New Yorker*, an exposure of trends in the fiction published by different editors could lead to a better understanding of why *New Yorker* short fiction, and perhaps short fiction more generally, has been dominated in recent years by particular genres or types of authors.

A second major unknown addressed in this article is the extent to which authors write fiction about characters similar to themselves, a question that has been raised time and again about authors and their work. Although literary criticism of individual pieces often addresses the degree to which a writer has drawn from personal experience, to our knowledge, an attempt has never been made to analyze this question with a large enough population to obtain results with statistical significance. With the sizeable amount of data available from 1990s *New Yorker* fiction, it is possible to determine how often *New Yorker* writers invent stories about characters who resemble themselves, as well as which of their own characteristics they most frequently project onto their protagonists and secondary characters. Essentially, this analysis will provide some insight into how autobiographical modern *New Yorker* short stories typically are.

The era selected for study is of interest because it is contemporary and because there were two different Executive Editors at the *New Yorker* during this time, as well as two different sets of Fiction Editors. In 1992, Tina Brown, the avant-garde editor of *Vanity Fair*, became the third Executive Editor of the *New Yorker*, shaking up its somewhat stilted image and drawing renewed attention to the periodical.[1] At the time, the *New Yorker*'s Fiction Department was under the stewardship of the magazine's Deputy Editor, Charles McGrath, a well-respected editor and writer who had been with the *New Yorker* since the early 1970s.[2] Two other senior Fiction Editors, Roger Angell and Daniel Menaker, largely shared McGrath's responsibilities and were essentially his equals in the Fiction Department's hierarchy. Both of these men had been members of the *New Yorker* staff for over 20 years by 1992,[3,4] and the three were considered to have fairly traditional tastes in fiction, in part due to the time they spent at the *New Yorker* while it was under the stewardship of conservative Executive Editor William Shawn. However, in early 1995, Bill Buford of the unconventional British literary magazine *Granta* was tapped to take over the *New Yorker* Fiction Department, and during his reign, the Department was controlled by one rather than multiple editors.[5] Buford's tenure continued beyond the end of the era studied in this article and weathered the 1998 transition to a new, less risqué Executive Editor, when Pulitzer Prize winning author and *New Yorker* staff writer David Remnick replaced Tina Brown.[6] The period analyzed in this article allows the significance of changes in both fiction and executive editors to be assessed and ends just before 11 September 2001, in order to avoid dealing with any alterations in *New Yorker* content that may have been a result of the terrorist attacks.

The data set of *New Yorker* fiction referenced in this article, which contains thirty-nine variables for each of the 442 stories examined, was collected by K.L. Milkman. The variables contained within this data set are listed in Table 1, and an explanation of the methodologies used in the compilation of this database is available upon request. Not every variable was available for every piece of fiction,

Table 1 *New Yorker* fiction data set variables: all stories in issues between 5 October 1992 and 17 September 2001

Story	Author	Protagonist	Secondary A	Secondary B
Date of publication	Name	Gender	Gender	Gender
Double/single issue	Gender	Country of origin	Country of origin	Country of origin
Title	Country of origin	Age range (decade)	Age range (decade)	Age range (decade)
Exact location	Home region(s)	Race	Race	Race
US state location	Age range (decade)	Occupation	Occupation	Occupation
Geographical region	Race	Sexuality	Sexuality	Sexuality
Topics (up to seven)	Sexuality	Religion	Religion	Religion
Language (if translated)	Religion			
Page length				
1st or 3rd person				

and missing variables were excluded from all analyses conducted.

In Section 2 of this article, we present an overview of past work on similar topics. In Section 3, we attempt to answer the question of whether or not changes in Executive Editor or Fiction Editor are associated with significant changes in the type of fiction published in the *New Yorker*. In Section 4, we examine whether or not *New Yorker* authors write fiction about characters who share their demographic characteristics. In Section 5, a summary of the main conclusions resulting from this article is presented.

2 Literature Review

Although a great deal of quantitative literary analysis has been conducted in the past, most notably by Burrows (1987), Craig (2004), and Hoover (1998), little if any analysis has been carried out using data sets compiled by a human reader. The majority of previous computational studies of literature have been conducted on data sets collected by computer programs, which are not sophisticated enough to determine the demographic characteristics of a protagonist, to classify secondary characters, or to detect story topics, among other things. Instead, past studies have focused on such subjects as authorship attribution, word frequency, and translation. It is the hope of the authors of this article that future computational studies of literature will include work that relies on human readers to create data sets containing attributes of fiction that are not detectable using computer codes.[7]

3 Study 1: Editorial Shifts and 1990s *New Yorker* Fiction

Even for those who have worked at the *New Yorker* for decades, it is difficult to understand or explain the impact of editorial changes on the magazine's fictional content. However, it is possible to analyze various components of *New Yorker* fiction to determine if and how they changed concurrently with the editorial power shifts that took place during the 1990s. Here we examine differences in a sample of 442 stories, 120 of which were published under Charles McGrath,[8] who served as the *New Yorker*'s highest ranking Fiction Editor between 5 October 1992 and 10 April 1995, and 322 of which were published under Bill Buford, who served as Fiction Editor of the *New Yorker* from 17 April 1995 through 17 September 2001. We also examine differences in the same sample of 442 stories published under successive Executive Editors: 275 under Tina Brown, who served as Executive Editor from 5 October 1992 until 3 August 1998, and 167 under David Remnick, who took over as Executive Editor on 10 August 1998 and was still serving in that capacity on 17 September 2001.

In order to evaluate what changes in *New Yorker* fiction were associated with different editorial shifts that took place during the 1990s, we examine the proportion of stories exhibiting various characteristics before and after editorial shakeups at the magazine. The discussion that follows presents our significant findings and refers readers to tables

in the Appendix for more detailed statistical evidence supporting each of our results.

3.1 Results

A number of components of *New Yorker* fiction changed significantly when there was a change in the magazine's Fiction Editor but remained unchanged when there was a switch in the *New Yorker*'s Executive Editor. We begin by discussing significant shifts in the proportions of stories with certain characteristics that accompanied Bill Buford's rise to the helm of the fiction department but were not associated with a change in the *New Yorker*'s Executive Editor. First, we find that the proportion of stories published with third-person narrators (as opposed to first-person narrators) increased significantly under Bill Buford but was not statistically different under Tina Brown and David Remnick (Fig. 1 and Appendix Table 2). Similarly, there was a significant increase in the percentage of fiction published by male authors when Bill Buford took over as Fiction Editor of the *New Yorker*, but the proportion of fiction by men published in the *New Yorker* was not significantly different under successive Executive Editors (Fig. 2 and Appendix Table 3). We also find that there was a significant increase in the variance of the page length of *New Yorker* fiction pieces published under Bill Buford as compared with pieces published under Charles McGrath, whereas there was no significant change in the variance in the length of fiction published under successive Executive Editors of the magazine (Fig. 3 and Appendix Table 4).

The Fiction Editor transition we studied also produced more shifts in the topics of *New Yorker* fiction than a change in the magazine's Executive Editor. We examine the proportion of stories focusing on each of twenty-eight topics (topics were selected for analysis in advance of the data collection stage) under each editor in our sample.[9] Bill[10] Buford's rise to control over the *New Yorker*'s fiction department was accompanied by seven significant shifts in the frequency of story topics, while the change in Executive Editor we studied was only accompanied by four significant shifts in the frequency of story topics (Figs 4 and 5, Appendix Tables 5 and 6).

In addition to examining significant shifts in the proportion of stories with certain characteristics that coincided with a change in the *New Yorker*'s Fiction Editor, we examine shifts in the distributions of story characteristics. We find that the distribution of story settings across the nine US Census regions pictured in Fig. 6 shifted significantly when there was a change in the *New Yorker*'s Fiction Editor but remained relatively unaltered under the different *New Yorker* Executive Editors in our sample. In particular, the proportion of stories
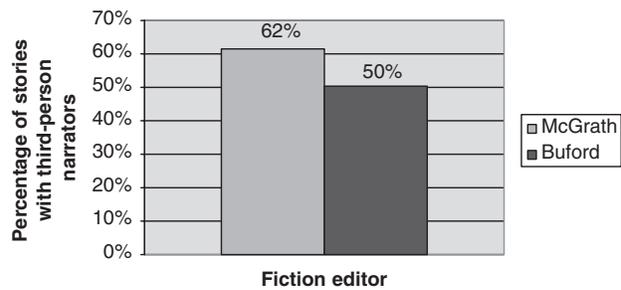

Fig. 1 Shift in percentage of stories with third-person narrators published under different Fiction Editors
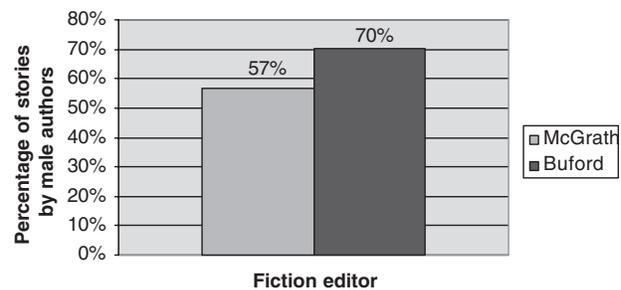

**Fig. 2** Shift in percentage of male authors published under different Fiction Editors
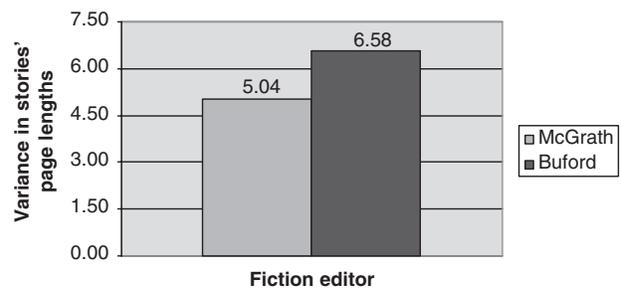
Fig. 3 Shift in variance of stories' page lengths under different Fiction Editors

set in the Middle Atlantic Region, which encompasses New York, increased dramatically when Bill Buford took over for Charles McGrath in mid-1995, but this statistic was not significantly affected by the mid-1998 change in Executive Editor (Fig. 7 and Appendix Table 7). The distribution of characters' religions also shifted significantly in the fiction published after Bill Buford replaced Charles McGrath at the helm of the *New Yorker's* Fiction Department, but there was no significant difference in the distribution of characters' religions in the fiction published under David Remnick and Tina Brown. Specifically, McGrath published fiction with a significantly higher percentage of Jewish characters than Buford (Fig. 8 and Appendix Table 8). We also find that the distributions of both protagonists' ages and characters' ages differed significantly under

successive stewards of the Fiction Department in the 1990s but not under successive Executive Editors. In particular, the proportion of protagonists and characters in their twenties increased significantly when Bill Buford replaced Charles McGrath, but these proportions remained relatively stationary when David Remnick replaced Tina Brown (Figs 9 and 10 and A1ppendix Table 9).

Four patterns in our data contradict the paradigm that Fiction Editors exert more control than Executive Editors over the fiction published at the *New Yorker*. First, we find that statistically significant shifts in the distributions of characters' nationalities occurred when David Remnick replaced Tina Brown, but there was not a significant shift in these underlying distributions when Bill Buford took over for Charles McGrath. However, the proportion
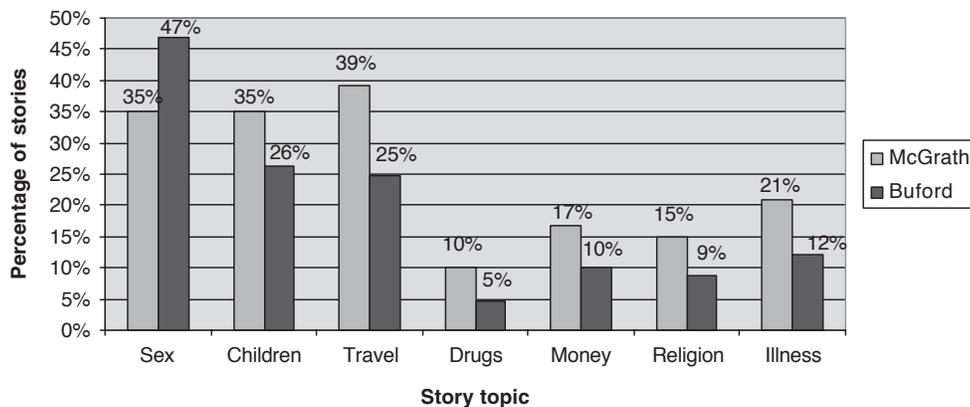

Fig. 4 Significant shifts in story topics under different Fiction Editors
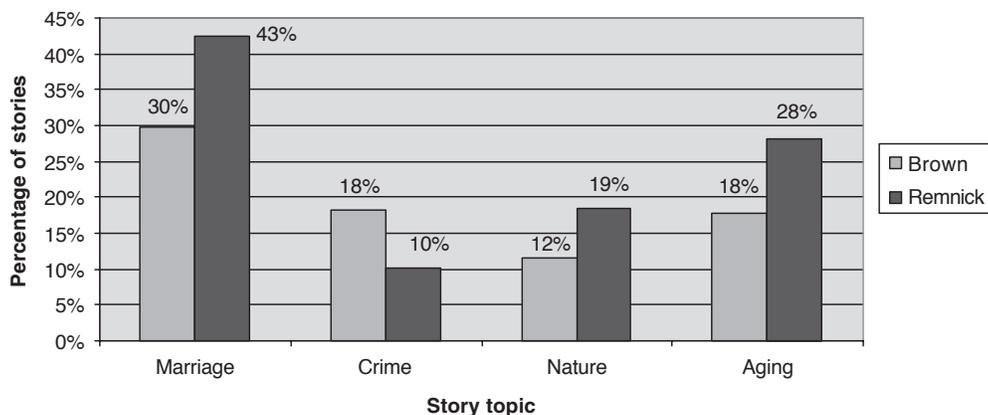

Fig. 5 Significant shifts in story topics under different Executive Editors
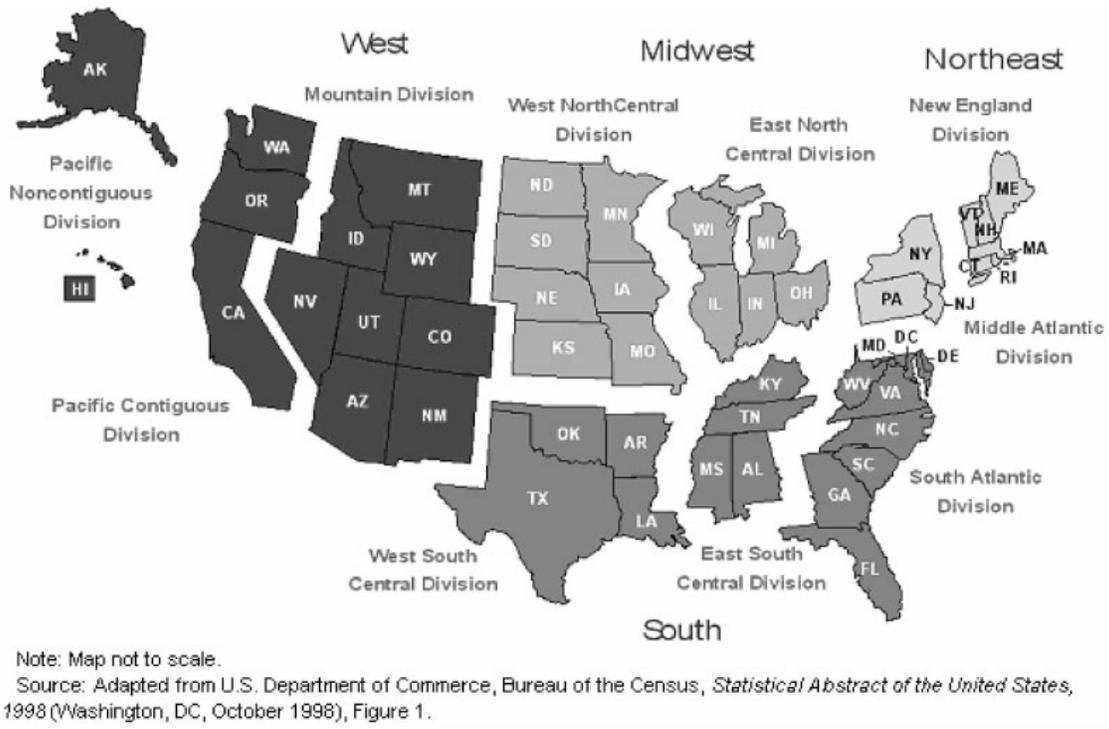
Fig. 6 Breakdown of nine US Census regions
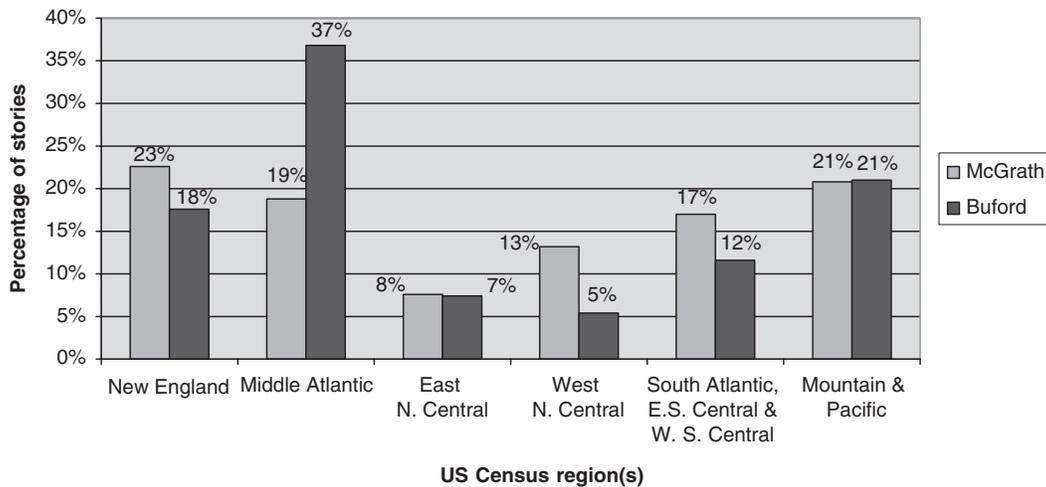
Fig. 7 Shift in distribution of stories set in each US Census region under different Fiction Editors

of characters from England decreased significantly when Bill Buford replaced Charles McGrath, but remained almost unchanged when David Remnick replaced Tina Brown (Figs 11 and 12 and Appendix Table 10).[11] We also find that there were significant shifts in the distributions of the races of characters and of authors when David Remnick replaced Tina Brown at the helm of the *New Yorker* but there were no shifts in these underlying distributions when Bill Buford took over for Charles McGrath (Figs 13 and 14 and Appendix Table 11). Finally, we observe a significant upward shift in the proportion
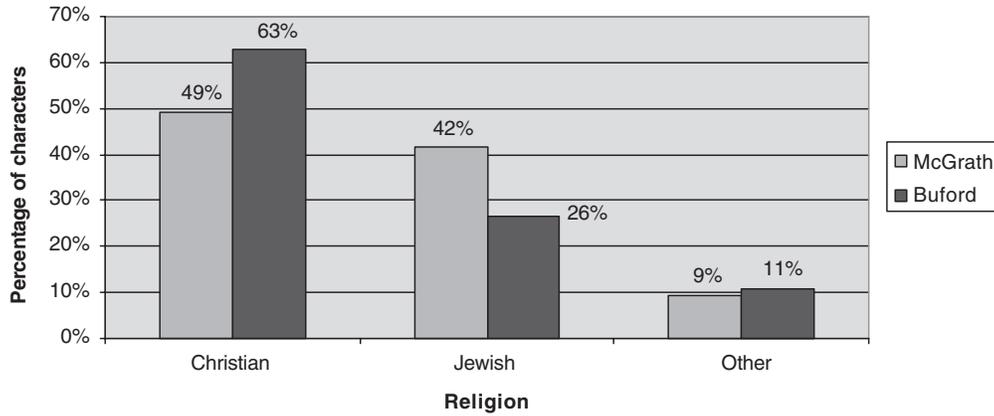
Fig. 8 Shift in distribution of characters' religions under different Fiction Editors
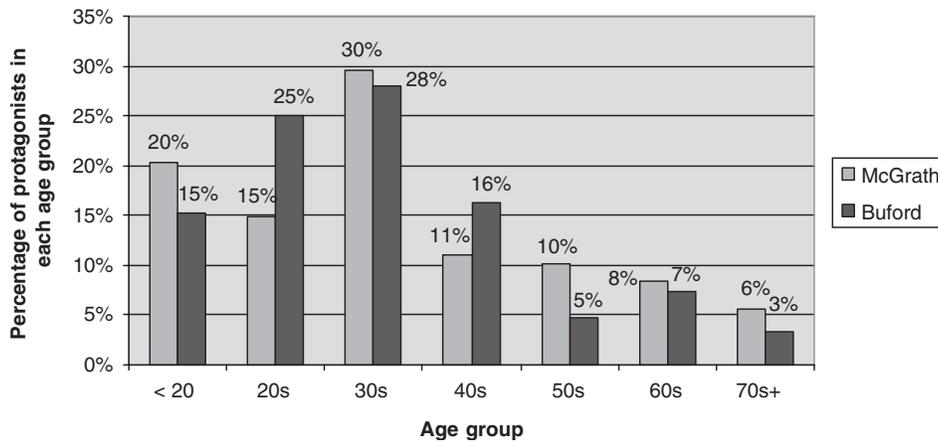

Fig. 9 Shift in distribution of protagonists' ages under different Fiction Editors
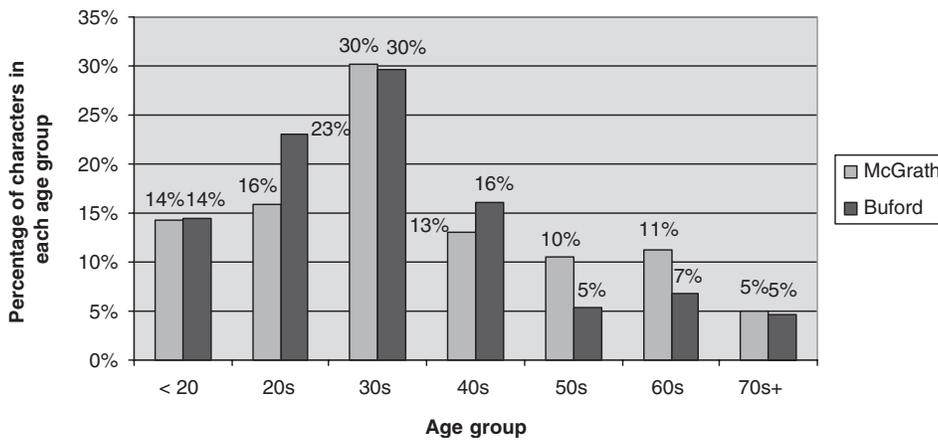

Fig. 10 Shift in distribution of characters' ages under different Fiction Editors
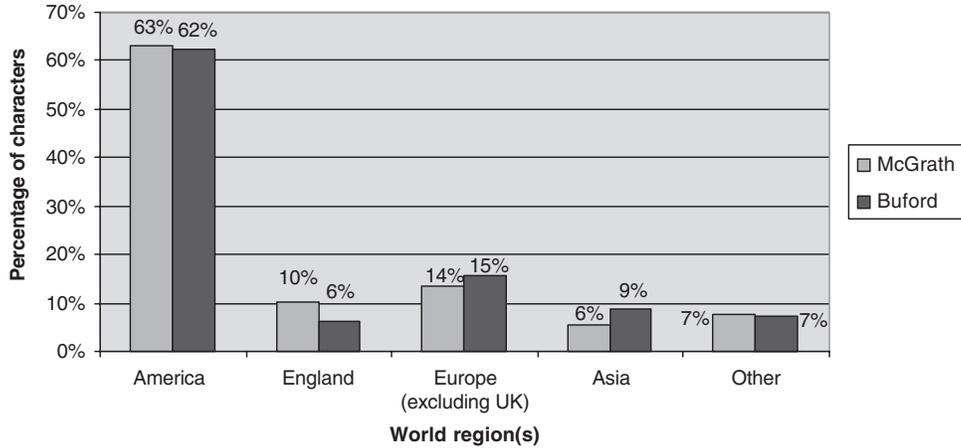
Fig. 11 Distributions of characters from each country of origin under different Fiction Editors
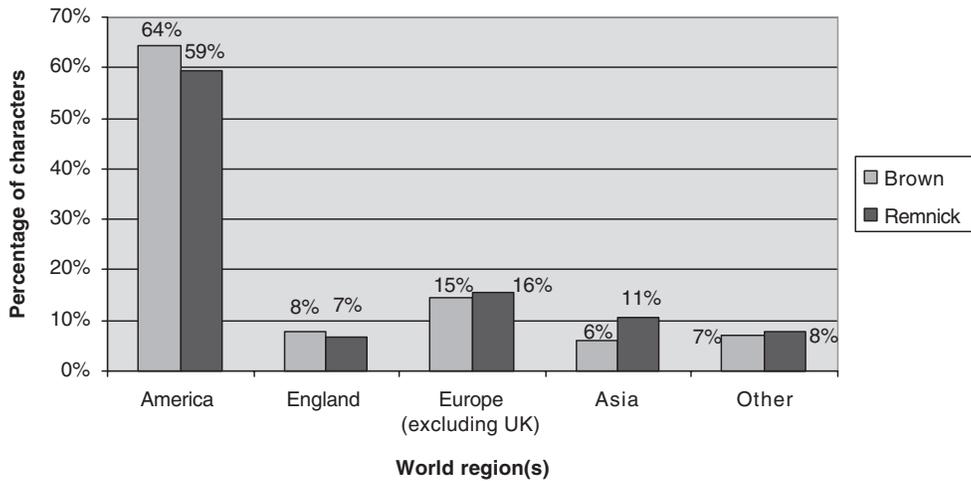
Fig. 12 Shift distribution of characters from each country of origin under different Executive Editors

of heterosexual characters in the fiction published after Remnick replaced Brown, while there was no change observed following the Fiction Department's transition from McGrath to Buford (Fig. 15 and Appendix Table 12).

3.2 Discussion

The statistical analyses presented in this section indicate that changes in the Fiction Editor of the *New Yorker* were associated with changes in various aspects of the magazine's fiction during the era studied and that changes in the *New Yorker*'s Executive Editor were associated with fewer

measurable changes in the fictional content of the magazine. Overall, we find that under Bill Buford:

- fewer stories were published with third-person narrators
- more stories were published by male authors
- there was more variance in stories' page length
- there were more stories about sex and fewer about children, travel, drugs, money, religion, and illness
- there was a shift in the geographic distribution of story settings, and more fiction was set in the Mid-Atlantic region
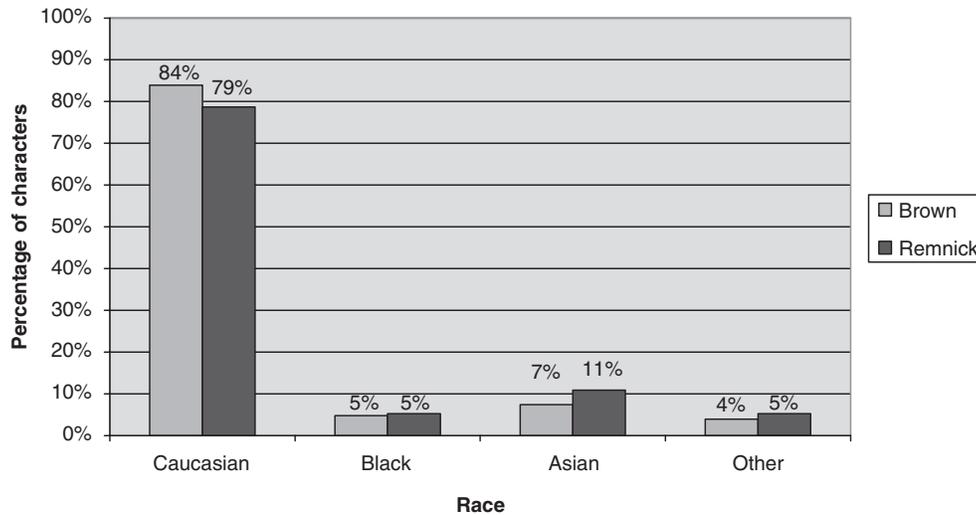
Fig. 13 Shift in distribution of characters of each race under different Executive Editors
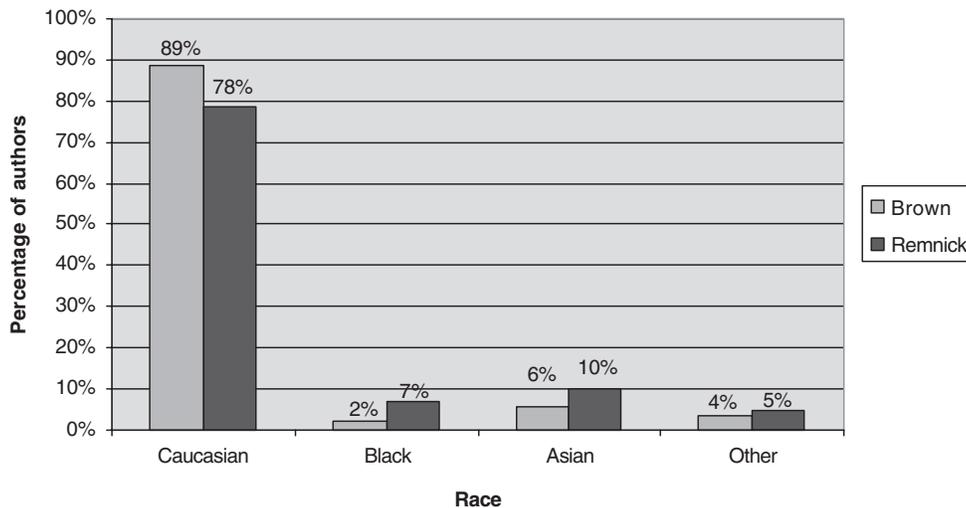
Fig. 14 Shift in distribution of authors of each race under different Executive Editors

- there was a shift in the distribution of characters' religions, and fewer characters in the fiction were Jewish
- the distribution of protagonists' and characters' ages shifted, and the proportion of protagonists and characters in their twenties increased
- there were fewer characters from England

Under David Remnick we find that:

- there were more stories about marriage, nature, and aging and fewer stories about crime

- there was a shift in the distribution of characters' countries of origin
- there were shifts in the distributions of characters' and authors' races
- more characters in the fiction were heterosexual

Additional tests we ran on the variables collected with respect to the question of how the fiction published under successive editors differed yielded no significant results but are available upon request. The results of the empirical analyses presented herein demonstrate that a quantitative
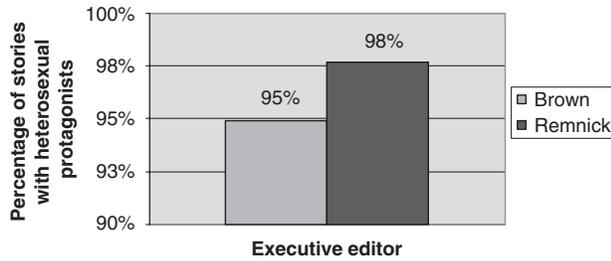
Fig. 15 Shift in percentage of stories with heterosexual protagonists published under different Executive Editors

study can help clarify the influence of different editors' preferences on a magazine's content, which might not otherwise be recognized or understood.

A significant question raised by the findings in this section is why a change in the New Yorker's Fiction Editor would be associated with such significant changes in the magazine's fiction. There are three plausible answers to this question: (1) the submissions the New Yorker received changed as a result of writers' knowledge of who was making selections at the magazine; (2) Charles McGrath had significantly different tastes in fiction than Bill Buford; and/or (3) our results are driven by general trends in short fiction in the 1990s. It is impossible to disentangle whether explanation (1) or (2) is responsible for our findings with the data available to us, but this could be a fruitful area for future research if the New Yorker would provide data about the characteristics of unpublished manuscripts submitted to its Fiction Department. We feel some confidence in ruling out explanation (3) for our results given that general trends in fiction would have manifested themselves in the analyses we conducted of the relationship between New Yorker fiction and a change in the New Yorker's Executive Editor, and they did not.

The findings from this study suggest that the New Yorker's Executive Editors have less influence over the magazine's fiction content than the New Yorker's Fiction Editors. However, it seems likely that these same Executive Editors have more significant influence over some other aspect of the New Yorker's content. It would be interesting to conduct a similar study of another magazine or even another division of the New Yorker to see how

editorial changes at various levels impact other types of magazine content. With respect to other literary magazines, a greater focus on fiction vis-a-vis other content might translate into greater shifts in fictional content as a result of a change in Executive Editor. It should also be acknowledged that the influence of Brown and Remnick on the New Yorker's fiction arguably filtered through their selections of fiction editors, and thus Brown's choice of Buford to replace McGrath could be viewed as her most significant way of exerting influence over the magazine's fiction.

Unfortunately, the size of the data set examined in this study was insufficient for addressing questions about the effects of a change in Fiction Editor under a single Executive Editor or the effects of a change in Executive Editor served by a single Fiction Editor. However, it would be interesting in a future study to see if Executive Editors perhaps have a more significant impact on the New Yorker's fiction when their influence is not disturbed by shifts in Fiction Editors. It would also be interesting to see if a change in Fiction Editor might have an even greater effect than detected in this study if successive Fiction Editors were observed during a time period when there was no change in Executive Editor.

4 Study 2: How Demographically Similar are the Authors of *New Yorker* Fiction and the Characters they Invent?

It has long been assumed that much of fiction contains elements of autobiography. However, it is difficult to imagine how one might quantify the extent to which a piece of fiction is autobiographical in order to test this hypothesis. There are some components of a work of fiction that can be measured, however, which may offer some insight into the question of how closely an author has modeled her protagonist and secondary characters on herself. These measurable components are indicators of whether an author shares various demographic characteristics with the protagonist
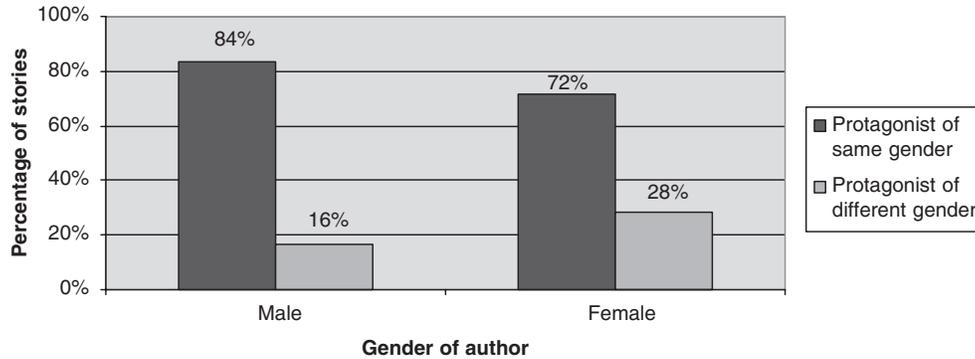
**Fig. 16** Relationship between gender of author and protagonist
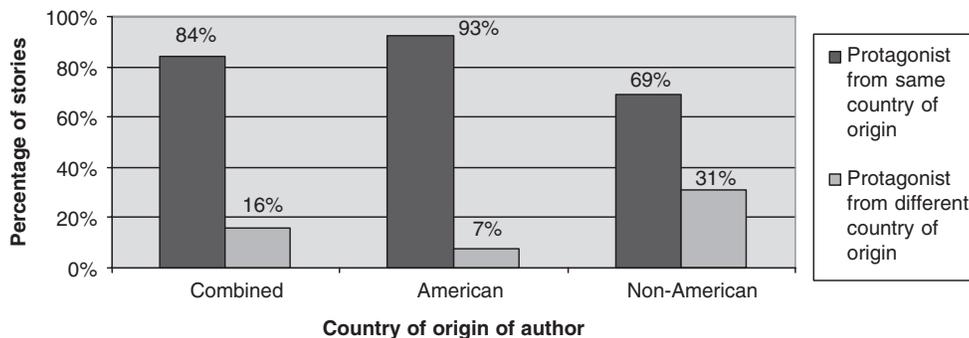
Fig. 17 Relationship between country of origin of author and protagonist

and secondary characters in a story. An examination of the relationship between author and character demographic characteristics at least begins to address the question of whether fictional characters are crafted, to some extent, in their authors' own image.

Using the sample of 442 *New Yorker* short stories whose characteristics we documented for this article, we examine the rate at which *New Yorker* authors write about characters who share various aspects of their demographic profile. In the discussion that follows we present our significant findings and refer readers to tables in the Appendix for more detailed statistical evidence supporting each of our results.

4.1 Results

We begin by discussing the demographic similarities between *New Yorker* authors and their protagonists. Consistent with the idea that authors write about what they know, we find that the vast majority of

New Yorker fiction authors write stories about protagonists of their gender (Fig. 16 and Appendix Table 13). Interestingly, male and female authors do this at differing levels—women write significantly more often than men about protagonists of the opposite gender. Similarly, we find that *New Yorker* authors project their native nationalities on to the protagonists they invent, the vast majority of the time (Fig. 17 and Appendix Table 14) with American authors writing about protagonists from their country of origin at a significantly higher rate than non-American authors.

In addition, we find that *New Yorker* authors share their race with their protagonists significantly more often than not, although again, authors from different racial groups vary in the frequency at which they write about protagonists who share their race (Fig. 18 and Appendix Table 15). There was only enough data to statistically examine the strength of the relationship between Caucasian authors and the protagonists in their stories,
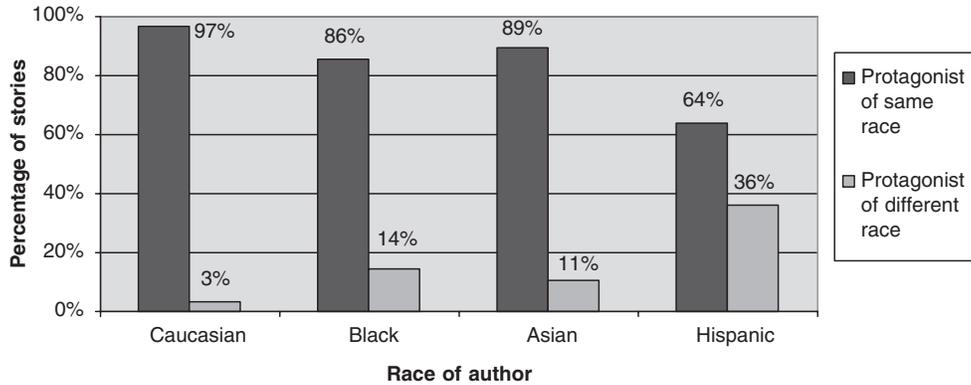
Fig. 18 Relationship between race of author and protagonist
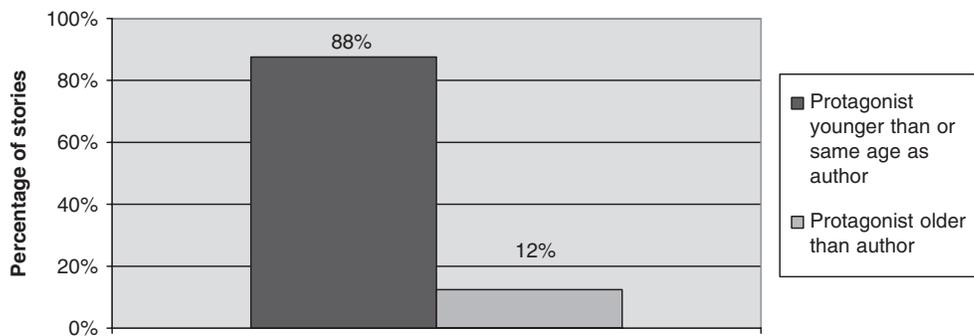
Fig. 19 Relationship between age of author and protagonist

and that relationship was extremely strong, with nearly all of the protagonists in stories written by Caucasian authors sharing their author's race. However, the other data collected indicates that there is also a strong relationship between the race of Black, Asian, and Hispanic authors and their protagonists. Finally, we find that the vast majority of *New Yorker* authors write about protagonists who are within or below their decade age range, or in other words, most protagonists are in an age group that their author has experienced (Fig. 19 and Appendix Table 16).[12] Slightly more than a fifth of protagonists in the fiction studied were within the precise decade age-range of the author of the story in which they appeared.

Next we turn to an examination of the demographic similarities between *New Yorker* authors and their secondary characters. We find that *New Yorker* authors are slightly less likely to write about

secondary characters of their gender than of the opposite gender (Fig. 20 and Appendix Table 17). At first glance, this seems at odds with our finding that *New Yorker* authors write about protagonists who share their gender the vast majority of the time. However, this pattern seems likely due to *New Yorker* authors' heavy reliance on heterosexual romantic plot lines,[13] which require the creation of characters of the opposite gender from protagonists.

We find that *New Yorker* authors share their nationality with the secondary characters in their stories significantly more often then not (Fig. 21 and Appendix Table 18). Again, this tendency is stronger among American authors who publish in the *New Yorker* than among authors of other nationalities. We also find that *New Yorker* authors project their race onto their secondary characters the vast majority of the time (Fig. 22 and
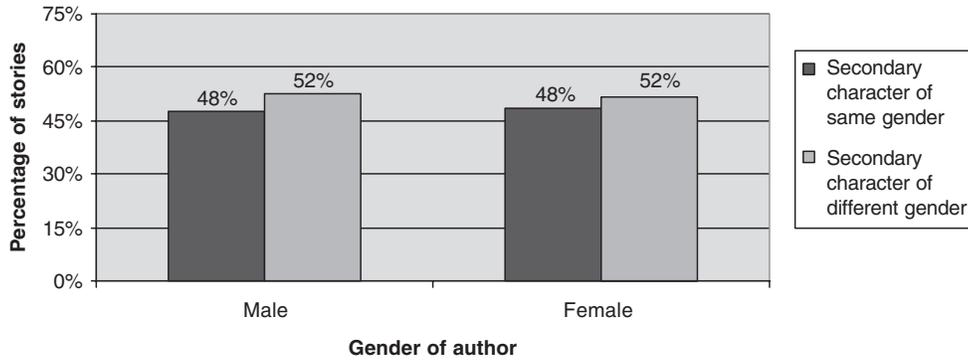
Fig. 20 Relationship between gender of author and secondary characters
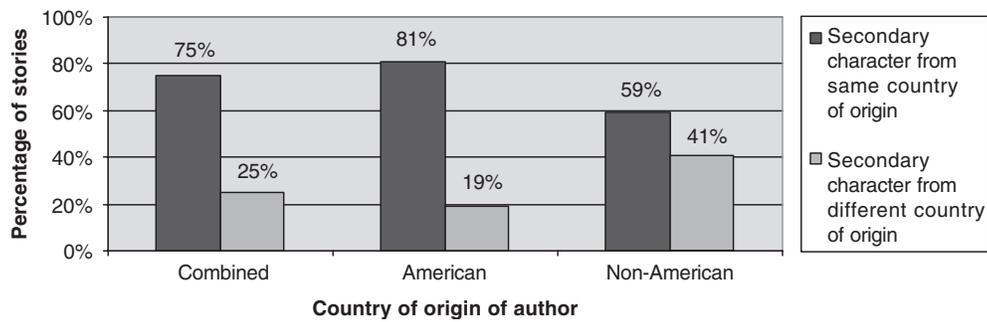
Fig. 21 Relationship between country of origin of author and secondary characters
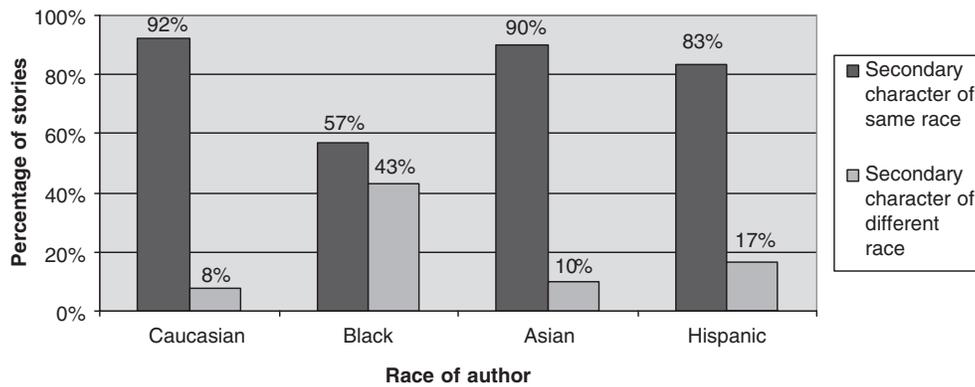
Fig. 22 Relationship between race of author and secondary characters

Appendix Table 19). Although very few stories in our sample were written by minorities, our data allows us to detect statistically significant differences in the rates at which both Caucasian authors and Black authors write about secondary characters of their race, and the limited data on Hispanic and Asian authors suggests that this pattern applies to them as well. Finally, we observe that the secondary characters in *New Yorker* fiction are almost always in or below the same age range as the authors of the stories in which they appear (Fig. 23 and Appendix Table 20). All of the demographic similarities between authors and secondary characters reported here are observed at lower
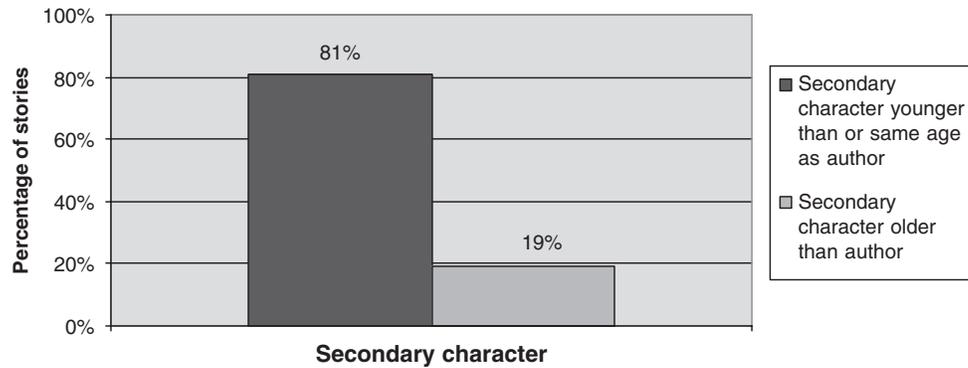
Fig. 23 Relationship between age of author and secondary chacters
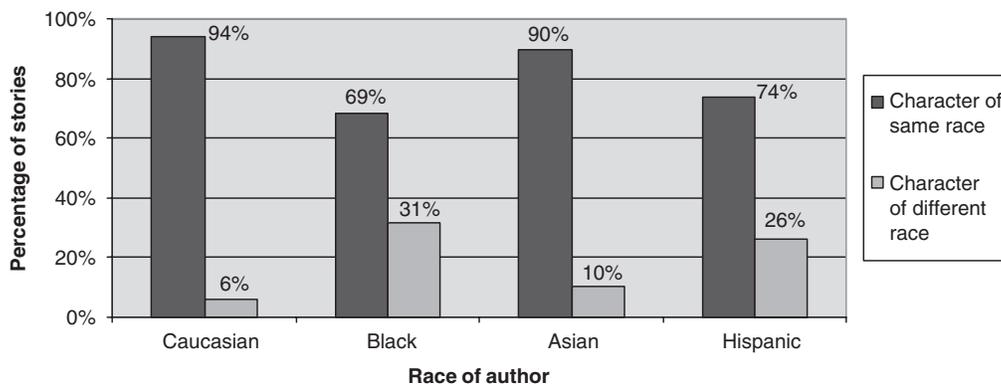
Fig. 24 Relationship between race of author and characters

frequencies than the similarities between authors and their protagonists, suggesting that authors may be more prone to create protagonists in their own image than secondary characters.

As mentioned earlier, with the exception of Caucasian authors, there was insufficient data to draw statistically supportable conclusions about the relationship between the race of *New Yorker* authors and their protagonists, and only for Caucasian and Black authors was it possible to draw such conclusions about the relationship between the race of authors and their secondary characters. Consequently, the two sets of data on protagonists and secondary characters were combined for analysis under the broader heading of 'characters'. This aggregated data allows us to conclude that Caucasian, Black, Hispanic, and Asian authors all write about characters who share their race significantly more often than not (Fig. 24 and Appendix Table 21). The rates at which Blacks

and Hispanic portray characters of their own race are extremely similar, yet considerably lower than those of Asian and Caucasian authors, who also appear to write about members of their own races at comparable frequencies.

Finally, we turn to an analysis of the settings selected for the stories in this study. We find that *New Yorker* authors not only write about characters with numerous similarities to themselves; they also write about settings that are familiar to them. Authors in our data set locate stories in their native world region significantly more often than not (Fig. 25 and Appendix Table 22),[14] and American authors write stories set in their home region of the world significantly more often than non-American authors.

4.2 Discussion

The data and statistical analyses presented in this section offer convincing support for the
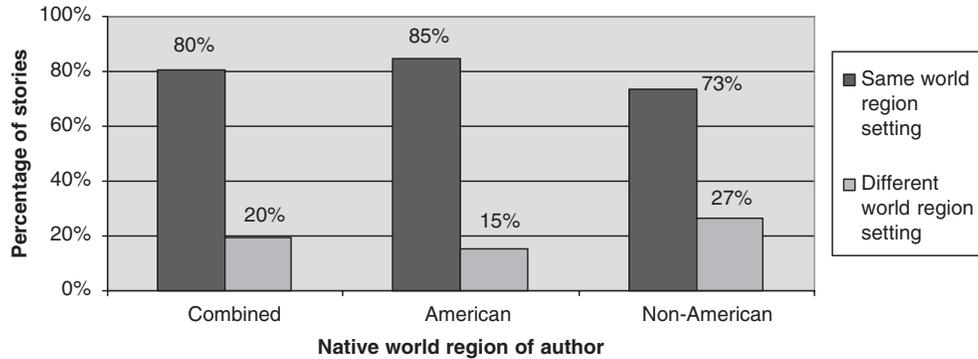
Fig. 25 Relationship between native world region of author and story setting

hypothesis that *New Yorker* fiction authors write more often than not about protagonists who share many of their demographic characteristics. The results also suggest that secondary characters have many demographic similarities to the authors of the stories in which they appear, albeit fewer than protagonists. It is not surprising that *New Yorker* authors of fiction write about what they know. However, to our knowledge, the commonly held hypothesis that authors of fiction write about characters who are similar to them has never been tested, making our study the first to provide a sound evidentiary basis for a widely held notion about fiction.

Our results indicate that *New Yorker* authors have a strong propensity to write about protagonists who share their gender, race, and nationality, and who are in or below their age range. Similarly, they write more often than not about secondary characters who share their race, their nationality, and are within or below their age range. Our results also suggest that *New Yorker* authors set their stories in locales where they have lived significantly more often than not. Perhaps more interesting than any of these observations, however, is our finding that *New Yorker* authors whose demographic characteristics make them minorities as far as the magazine's publishing roster is concerned (i.e. females, non-Whites, and non-Americans) are less likely than others to write about protagonists and secondary characters who share their demographic characteristics. Given the data available to us, it is not possible to determine if this is

due to a selection bias at the *New Yorker*, the choices made by authors about what fiction to submit to the *New Yorker*, or general tendencies among minority authors to conform to fictional norms. It would be interesting if a future study were able to disentangle these alternative explanations for our results.

In addition to this potentially fruitful area for future research, our findings suggest a number of additional topics for future study. It seems that further analysis of the relationship between the race, sexuality, and religion of authors and their characters could yield interesting results. It might also be worthwhile to look at the same aspects of fiction studied in this article within the context of novels and other sets of short fiction besides the stories published by the *New Yorker* in the 1990s in order to understand whether or not the conclusions drawn in this article apply to fictional literature as a whole, short fiction as a whole, or some other, narrower literary grouping.

5 Conclusions

In this article, an extensive quantitative analysis of two questions with regard to the 442 pieces of fiction published between 5 October 1992 and 17 September 2001 in the *New Yorker* magazine yielded several interesting results. The analysis presented in Section 3 examined shifts in the *New Yorker*'s fictional content that were associated with a change in the magazine's Fiction Editor as well as shifts associated with a change in the magazine's

Executive Editor. We find that a change in the *New Yorker's* Fiction Editor is associated with numerous significant shifts in the magazine's fictional content. Specifically, we find that a change in the *New Yorker's* Fiction Editor is associated with shifts in such variables as narrative voice, the regional distribution of stories set in the US, the gender of authors, the distribution of characters' religions, the variance in stories' lengths, the age distributions of protagonists and characters as a whole, the frequency of seven story topics, and the proportion of British characters in *New Yorker* short stories. Quantitative analyses reported on in Section 3 also led to the conclusion that the changes in *New Yorker* fiction associated with Bill Buford's replacement of Charles McGrath as the *New Yorker's* Fiction Editor were more pronounced than the changes associated with David Remnick's replacement of Tina Brown as the magazine's Executive Editor. The shift in Executive Editor studied in this article was only accompanied by changes in the distribution of characters' countries of origin, authors' and characters' racial distributions, characters' sexuality, and four shifts in topic frequency.

As reported in Section 4 of this article, quantitative analysis demonstrated that, in the vast majority of cases, authors of *New Yorker* fiction write about protagonists who share their race, gender, and nationality and who are within or below their age range. These findings are consistent with the widely held notion that much of fiction is autobiographical. We also find that, to a slightly lesser degree, *New Yorker* authors invent secondary characters who share their race and nationality and who are within or below their age range. However, most likely in deference to heterosexual plot considerations, authors write about secondary characters who, slightly more than half the time, do not share their gender. This finding suggests that *New Yorker* secondary characters are crafted to resemble members of the demographic group that authors are most familiar with, but that secondary characters are less likely than protagonists to be autobiographical representations of an author. Finally, we conclude that the majority of the time *New Yorker* authors write stories set in their home region(s). All of these findings point to the conclusion that the quantifiable components of short stories in the *New Yorker* are largely based on what an author knows and that characters are, to a great extent, demographically similar to the authors who invent them.

In addition to presenting specific findings, this article demonstrates several ways in which statistical analysis can successfully be employed to address questions about literature. Although quantitative methods have been used before to examine trends in literature, the traditional approach has been to examine questions that do not require content analysis by a human reader. This article uses variables derived both from reading stories and from outside sources to draw conclusions about literature. It is the hope of the authors that future research will further examine the questions raised in this article by looking at other types of literature and asking an aneven broader range of questions about the impact of editors on fiction and the similarities between authors and the characters they invent. However, care should always be taken in such studies to select a data set with objective variables in order to avoid the danger of computing statistics influenced by the interpretation of a human reader.

References

Burrows, J. F. (1987). *Computation Into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Clarendon Press.

Contemporary Authors. (2004). *Literature Research Center*. Gale Group Databases. [Internet] Princeton, NJ: Princeton University Libraries. <http://galenet.galegroup.com/servlet/LitRC?locID=princetonu> (accessed 25 January 2004).

Craig, H. (2004). A Stylistic Analysis and Authorship Studies. In Schreibman, S., Seimens, R., and Unsworth, J. (eds), *A Companion to the Digital Humanities*. Oxford: Blackwell Publishing.

Fiction. (June 1992–September 2001). *New Yorker*.

Literature Online. (2004). [*Internet*] *ProQuest Information and Learning Company*. <http://lion.chadwyck.com/> (accessed 25 January 2004).

Hoover, D. L. (1998). *Making Use of Statistical Measures of Style*, *MLA Convention*, San Francisco, 28 December 1998.

Notes

1 Gross, M. (1992). Tina's turn: The New Yorker's head transplant, *New York*, 20–31.
2 Anderson, C. (1998). Flexibility is key in publishing careers, says NYT book review editor, *The Harvard University Gazette*. Cambridge, MA: Harvard University. Online. 6 August.
3 Angell, R. (2004). *Contemporary Authors, Literature Research Center*, Gale Group Databases, Princeton NJ: Princeton University Libraries. Online. 3 January.
4 Untitled, http://www.geocities.com/nhsclassof59/menaker-Random-Hs-bio.html. Online. 3 January 2004.
5 *New Yorker* Names a Fiction Editor. (1994). *New York Times*, C22.
6 Granatstein, L. (1998). Remnick: A new *New Yorker*, *Mediaweek*, 22–23.
7 It should be noted that whenever data are collected by a human reader for a study, an element of subjectivity is introduced. This is the danger of relying upon people to create data sets about fiction rather than computers. However, it is the belief of the authors of this article that if the set of variables being cataloged are defined carefully enough to be largely objective and if the data set collected is large enough, this danger can be limited sufficiently to cause little if any variation in results from one experimenter to the next.
8 Although Roger Angell and Daniel Menaker exerted significant control over the *New Yorker*'s Fiction Department during the early 1990s, from this point on Charles McGrath will be referred to as the Fiction Editor who preceded Bill Buford because McGrath held the highest title of the three.
9 It should be noted that many of the twenty-eight topics coded as present or absent from a story are highly correlated (e.g., marriage and family).
10 Census Regions and Divisions, (2003). Energy Information Administration, 28 April 2003 <http://www.eia.doe.gov/cneaf/electricity/page/co2_report/fig1.html>.
11 This surprising decrease in the number of British characters in *New Yorker* fiction printed under Buford—a native of England—may be explained by the Editor's jaded attitude towards traditional British writing. [Bill Buford, Telephone interview with K.L Milkman (18 March 2004)].
12 The average age of the *New Yorker* authors in this sample was 49.8 years. Thus, the fact that authors were found to write the vast majority of the time about protagonists in or, to an even greater degree, below their age range could be a coincidence based on the fact that the majority of characters, like the majority of Americans (73% according 2000 Census figures), were under the age 50 as of 2000.
13 Around 96% of the characters in *New Yorker* fiction during the era examined were heterosexuals.
14 In this article, the world regions classified are: the United States, Western Europe, Eastern Europe, Asia, Middle East, Africa, North America (excluding the United States), and South America.

Appendix of Statistical Results

Table 2 Tests of shifts in percentage of stories with third-person narrators under different editors

Null hypothesis	Sample size	Test	*P*-value
Proportion of stories with third-person narrators under McGrath (62%) ≤ proportion of stories with third person narrators under Buford (50%)	McGrath: 120 Buford: 329	Proportion	0.023
Proportion of stories with third-person narrators under Brown (53%) ≥ proportion of stories with third person narrators under Remnick (55%)	Brown: 283 Remnick: 166	Proportion	0.364

Table 3 Tests of shifts in percentage of stories with male authors under different editors

Null hypothesis	Sample size	Test	*P*-value
Proportion of stories with male authors under McGrath (57%) ≥ proportion of stories with male authors under Buford (70%)	McGrath: 115 Buford: 321	Proportion	0.005
Proportion of stories with male authors under Brown (67%) ≤ proportion of stories with male authors under Remnick (66%)	Brown: 270 Remnick: 166	Proportion	0.476

Table 4 Tests of shifts in variance of story length under different editors

Null hypothesis	Sample size	Test	*P*-value
The variances in *New Yorker* fiction story length under McGrath (5.04) ≥ the variances under Buford (6.58)	McGrath: 120 Buford: 321	*F*-variance	0.045
The variances in *New Yorker* fiction story length under Brown (5.78) ≥ the variances under Remnick (6.74)	Brown: 275 Remnick: 166	*F*-variance	0.1301

Table 5 Proportion tests of shifts in frequency of story topics under different fiction editors

Topic	% of stories on topic printed by McGrath (Sample size:120)	% stories on topic printed by Buford (Sample size: 322)	Null hypothesis	*P*-value
Children	$P_{M,children} = 36\%$	$P_{B,children} = 26\%$	$P_{M,children} \leq P_{B,children}$	0.034
Drugs	$P_{M,drugs} = 10\%$	$P_{B,drugs} = 5\%$	$P_{M,drugs} \leq P_{B,drugs}$	0.031
Illness	$P_{M,illness} = 21\%$	$P_{B,illness} = 12\%$	$P_{M,illness} \leq P_{B,illness}$	0.015
Money	$P_{M,money} = 17\%$	$P_{B,money} = 10\%$	$P_{M,money} \leq P_{B,money}$	0.037
Religion	$P_{M,religion} = 15\%$	$P_{B,religion} = 9\%$	$P_{M,religion} \leq P_{B,religion}$	0.040
Sex	$P_{M,sex} = 35\%$	$P_{B,sex} = 47\%$	$P_{M,sex} \geq P_{B,sex}$	0.016
Travel	$P_{M,travel} = 39\%$	$P_{B,travel} = 25\%$	$P_{M,travel} \leq P_{B,travel}$	0.002
Adultery	$P_{M,adultery} = 12\%$	$P_{B,adultery} = 17\%$	$P_{M,adultery} \geq P_{B,adultery}$	0.120
Aging	$P_{M,aging} = 20\%$	$P_{B,aging} = 22\%$	$P_{M,aging} \geq P_{B,aging}$	0.368
Alcohol	$P_{M,alcohol} = 9\%$	$P_{B,alcohol} = 8\%$	$P_{M,alcohol} \leq P_{B,alcohol}$	0.388
Art	$P_{M,art} = 17\%$	$P_{B,art} = 13\%$	$P_{M,art} \leq P_{B,art}$	0.181
Betrayal	$P_{M,betrayal} = 13\%$	$P_{B,betrayal} = 20\%$	$P_{M,betrayal} \geq P_{B,betrayal}$	0.120
Crime	$P_{M,crime} = 20\%$	$P_{B,crime} = 13\%$	$P_{M,crime} \leq P_{B,crime}$	0.057
Death	$P_{M,death} = 38\%$	$P_{B,death} = 39\%$	$P_{M,death} \geq P_{B,death}$	0.470

(continued)

Table 5 Continued

Topic	% of stories on topic printed by McGrath (Sample size:120)	% stories on topic printed by Buford (Sample size: 322)	Null hypothesis	*P*-value
Divorce	$P_{M,divorce} = 8\%$	$P_{B,divorce} = 6\%$	$P_{M,divorce} \leq P_{B,divorce}$	0.241
Education	$P_{M,education} = 14\%$	$P_{B,education} = 14\%$	$P_{M,educate} \geq P_{B,educate}$	0.487
Family	$P_{M,family} = 37\%$	$P_{B,family} = 42\%$	$P_{M,family} \geq P_{B,family}$	0.202
History	$P_{M,history} = 6\%$	$P_{B,history} = 4\%$	$P_{M,history} \leq P_{B,history}$	0.240
Homosexuality	$P_{M,homosexuality} = 4\%$	$P_{B,homosexuality} = 4\%$	$P_{M,homosexuality} \geq P_{B,homosexuality}$	0.467
Love	$P_{M,love} = 28\%$	$P_{B,love} = 27\%$	$P_{M,love} \leq P_{B,love}$	0.4857
Marriage	$P_{M,marriage} = 28\%$	$P_{B,marriage} = 37\%$	$P_{M,marriage} \geq P_{B,marriage}$	0.057
Nature	$P_{M,nature} = 17\%$	$P_{B,nature} = 13\%$	$P_{M,nature} \leq P_{B,nature}$	0.232
Politics	$P_{M,politics} = 7\%$	$P_{B,politics} = 4\%$	$P_{M,politics} \leq P_{B,politics}$	0.183
Race	$P_{M,race} = 9\%$	$P_{B,race} = 10\%$	$P_{M,race} \geq P_{B,race}$	0.442
Relationships	$P_{M,relationships} = 58\%$	$P_{B,relationships} = 55\%$	$P_{M,relations} \leq P_{B,relations}$	0.335
Violence	$P_{M,violence} = 25\%$	$P_{B,violence} = 21\%$	$P_{M,violence} \leq P_{B,violence}$	0.228
War	$P_{M,war} = 7\%$	$P_{B,war} = 8\%$	$P_{M,war} \geq P_{B,war}$	0.426
Work	$P_{M,work} = 28\%$	$P_{B,work} = 22\%$	$P_{M,work} \leq P_{B,work}$	0.105

Table 6 Proportion tests of shifts in frequency of story topics under different executive editors

Topic	% of stories on topic printed by Brown (Sample size: 275)	% stories on topic printed by Remnick (Sample size: 167)	Null hypothesis	*P*-value
Aging	$P_{T,aging} = 18\%$	$P_{R,aging} = 28\%$	$P_{T,aging} \geq P_{R,aging}$	0.007
Crime	$P_{T,crime} = 18\%$	$P_{R,crime} = 10\%$	$P_{T,crime} \leq P_{R,crime}$	0.016
Marriage	$P_{T,marriage} = 30\%$	$P_{R,marriage} = 43\%$	$P_{T,marriage} \geq P_{R,marriage}$	0.004
Nature	$P_{T,nature} = 12\%$	$P_{R,nature} = 19\%$	$P_{T,nature} \geq P_{R,nature}$	0.030
Adultery	$P_{T,adultery} = 14\%$	$P_{R,adultery} = 17\%$	$P_{T,adultery} \geq P_{R,adultery}$	0.223
Alcohol	$P_{T,alcohol} = 8\%$	$P_{R,alcohol} = 8\%$	$P_{T,alcohol} \geq P_{R,alcohol}$	0.443
Art	$P_{T,art} = 15\%$	$P_{R,art} = 12\%$	$P_{T,art} \leq P_{R,art}$	0.204
Betrayal	$P_{T,betrayal} = 18\%$	$P_{R,betrayal} = 19\%$	$P_{T,betrayal} \geq P_{R,betrayal}$	0.410
Children	$P_{T,children} = 28\%$	$P_{R,children} = 31\%$	$P_{T,children} \geq P_{R,children}$	0.249
Death	$P_{T,death} = 35\%$	$P_{R,death} = 42\%$	$P_{T,death} \geq P_{R,death}$	0.098
Divorce	$P_{T,divorce} = 8\%$	$P_{R,divorce} = 4\%$	$P_{T,divorce} \leq P_{R,divorce}$	0.085
Drugs	$P_{T,drugs} = 7\%$	$P_{R,drugs} = 5\%$	$P_{T,drugs} \leq P_{R,drugs}$	0.387
Education	$P_{T,education} = 14\%$	$P_{R,education} = 14\%$	$P_{T,educate} \geq P_{R,educate}$	0.478
Family	$P_{T,family} = 37\%$	$P_{R,family} = 45\%$	$P_{T,family} \geq P_{R,family}$	0.074
History	$P_{T,history} = 4\%$	$P_{R,history} = 4\%$	$P_{T,history} \leq P_{R,history}$	0.466
Homosexuality	$P_{T,homosexuality} = 5\%$	$P_{R,homosexuality} = 4\%$	$P_{T,homosexuality} \leq P_{R,homosexuality}$	0.371
Illness	$P_{T,illness} = 16\%$	$P_{R,illness} = 13\%$	$P_{T,illness} \leq P_{R,illness}$	0.227
Love	$P_{T,love} = 27\%$	$P_{R,love} = 28\%$	$P_{T,love} \geq P_{R,love}$	0.475
Money	$P_{T,money} = 13\%$	$P_{R,money} = 10\%$	$P_{T,money} \leq P_{R,money}$	0.169
Politics	$P_{T,politics} = 5\%$	$P_{R,politics} = 4\%$	$P_{T,politics} \leq P_{R,politics}$	0.421
Race	$P_{T,race} = 8\%$	$P_{R,race} = 12\%$	$P_{T,race} \geq P_{R,race}$	0.112
Relationships	$P_{T,relationships} = 58\%$	$P_{R,relationships} = 51\%$	$P_{T,relations} \leq P_{R,relations}$	0.082
Religion	$P_{T,religion} = 12\%$	$P_{R,religion} = 8\%$	$P_{T,religion} \leq P_{R,religion}$	0.106
Sex	$P_{T,sex} = 41\%$	$P_{R,sex} = 48\%$	$P_{T,sex} \geq P_{R,sex}$	0.097
Travel	$P_{T,travel} = 29\%$	$P_{R,travel} = 28\%$	$P_{T,travel} \leq P_{R,travel}$	0.374
Violence	$P_{T,violence} = 20\%$	$P_{R,violence} = 26\%$	$P_{T,violence} \geq P_{R,violence}$	0.063
War	$P_{T,war} = 8\%$	$P_{R,war} = 7\%$	$P_{T,war} \leq P_{R,war}$	0.431
Work	$P_{T,work} = 23\%$	$P_{R,work} = 25\%$	$P_{T,work} \geq P_{R,work}$	0.337

Table 7 Tests of shifts in distribution of stories set in each US region under different editors

Null hypothesis	Sample size	Test	*P*-value
The distributions of story settings by US Region in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 53 Buford: 147	χ^2 GOF	0.028
Proportion of stories set in Middle Atlantic under McGrath (19%) ≥ proportion of stories set in Middle Atlantic under Buford (37%)	McGrath: 53 Buford: 147	Proportion	0.013
The distributions of story settings by US Region in *New Yorker* fiction are the same under Brown and Remnick	Brown: 124 Remnick: 76	χ^2 GOF	0.880
Proportion of stories set in Middle Atlantic under Brown (31%) ≥ proportion of stories set in Middle Atlantic under Remnick (33%)	Brown: 124 Remnick: 76	Proportion	0.478

Table 8 Tests of shifts in distribution of stories about characters of each religion under different editors

Null hypothesis	Sample size	Test	*P*-value
The distributions of characters' religions in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 53 Buford: 102	χ^2 GOF	0.041
Proportion of Jewish characters under McGrath (42%) ≤ proportion of Jewish characters under Buford (26%)	McGrath: 53 Buford: 102	Proportion	0.042
The distributions of characters' religions in *New Yorker* fiction are the same under Brown and Remnick	Brown: 99 Remnick: 56	χ^2 GOF	0.191
Proportion of Jewish characters under Brown (27%) ≥ Proportion of Jewish characters under Remnick (34%)	Brown: 99 Remnick: 56	Proportion	0.214

Table 9 Tests of shifts in distribution of ages of protagonists and characters under different editors

Null hypothesis	Sample size	Test	*P*-value
The distributions of protagonists' ages in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 108 Buford: 296	χ^2 GOF	0.011
Proportion of protagonists in their twenties under McGrath (15%) ≥ proportion of protagonists in their twenties under Buford (25%)	McGrath: 108 Buford: 296	Proportion	0.021
The distributions of characters' ages in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 239 Buford: 666	χ^2 GOF	$< 10^{-3}$
Proportion of characters in their twenties under McGrath (16%) ≥ proportion of characters in their twenties under Buford (23%)	McGrath: 239 Buford: 666	Proportion	0.014
The distributions of protagonists' ages in *New Yorker* fiction are the same under Brown and Remnick	Brown: 247 Remnick: 157	χ^2 GOF	0.702
Proportion of protagonists in their twenties under Brown (20%) ≥ proportion of protagonists in their twenties under Remnick (25%)	Brown: 247 Remnick: 157	Proportion	0.134
The distributions of characters' ages in *New Yorker* fiction are the same under Brown and Remnick	Brown: 545 Remnick: 360	χ^2 GOF	0.547
Proportion of characters in their twenties under Brown (20%) ≥ proportion of characters in their twenties under Remnick (23%)	Brown: 545 Remnick: 360	Proportion	0.179

Table 10 Tests of shifts in distribution of characters' countries of origin under different editors

Null hypothesis	Sample size	Test	*P*-value
The distributions of characters' countries of origin in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 214 Buford: 542	χ^2 GOF	0.0796
Proportion of characters from England under McGrath (10%) ≤ proportion of characters from England under Buford (6%)	McGrath: 214 Buford: 542	Proportion	0.041
The distributions of characters' countries of origin in *New Yorker* fiction are the same under Brown and Remnick	Brown: 295 Remnick: 461	χ^2 GOF	0.0211
Proportion of characters from England under Brown (8%) ≤ proportion of characters from England under Remnick (7%)	Brown: 295 Remnick: 461	Proportion	0.350

Table 11 Tests of shifts in distribution of characters' and authors' racial characteristics under different editors

Null hypothesis	Sample size	Test	*P*-value
The distributions of characters' races in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 256 Buford: 656	χ^2 GOF	0.1444
The distributions of authors' races in *New Yorker* fiction are the same under McGrath and Buford	McGrath: 101 Buford: 314	χ^2 GOF	0.3700
The distributions of characters' races in *New Yorker* fiction are the same under Brown and Remnick	Brown: 562 Remnick: 350	χ^2 GOF	0.0315
The distributions of authors' races in *New Yorker* fiction are the same under Brown and Remnick	Brown: 253 Remnick: 162	χ^2 GOF	0.0001

Table 12 Tests of shifts in proportion of stories about heterosexual characters under different editors

Null hypothesis	Sample size	Test	*P*-value
Proportion of heterosexual characters under McGrath (96%) ≤ proportion of heterosexual characters under Buford (96%)	McGrath: 221 Buford: 614	Proportion	0.457
Proportion of heterosexual characters under Brown (95%) ≥ proportion of heterosexual characters under Remnick (98%)	Brown: 493 Remnick: 342	Proportion	0.035

Table 13 Tests of relationship between gender of author and gender of protagonist

Null hypothesis	Sample size	Test	*P*-value
Proportion of male authors who write about male protagonists (84%) ≤ proportion of male authors who write about female protagonists (16%)	228 stories	Binomial	$<10^{-3}$
Proportion of female authors who write about female protagonists (72%) ≤ proportion of female authors who write about male protagonists (28%)	144 stories	Binomial	$<10^{-3}$
Proportion of male authors who write about male protagonists (84%) ≤ proportion of female authors who write about female protagonists (72%)	Male: 228 Female: 144	Proportion	0.008

Table 14 Tests of relationship between country of origin of author and country of origin of protagonist

Null hypothesis	Sample size	Test	P-value
Proportion of authors who write about protagonists from their country of origin (84%) $\leq$ proportion of authors who write about protagonists who do not share their country of origin (16%)	375 stories	Binomial	$<10^{-3}$
Proportion of American authors who write about protagonists from America (93%) $\leq$ proportion of American authors who write about protagonists who do not share their country of origin (7%)	243 stories	Binomial	$<10^{-3}$
Proportion of non-American authors who write about protagonists from their country of origin (69%) $\leq$ proportion of non-American authors who write about protagonists who do not share their country of origin (31%)	132 stories	Binomial	$<10^{-3}$
Proportion of American authors who write about protagonists from their country of origin (93%) $\leq$ proportion of non-American authors who write about protagonists from their country of origin (69%)	American: 243 Other: 132	Proportion	$<10^{-3}$

Table 15 Tests of relationship between race of author and race of protagonist

Null hypothesis	Sample size	Test	P-value
Proportion of Caucasian protagonists in stories by Caucasians (97%) $\leq$ proportion of non-Caucasian protagonists in stories by Caucasians (3%)	308 stories	Binomial	$<10^{-3}$
Proportion of black protagonists in stories by blacks (86%) $\leq$ proportion of non-black protagonists in stories by blacks (14%)	14 stories	Too little data to use Binomial	NA
Proportion of Asian protagonists in stories by Asians (89%) $\leq$ proportion of non-Asian protagonists in stories by Asians (11%)	28 stories	Too little data to use Binomial	NA
Proportion of Hispanic protagonists in stories by Hispanics (64%) $\leq$ proportion of non-Hispanic protagonists in stories by Hispanics (36%)	11 stories	Too little data to use Binomial	NA

Table 16 Test of relationship between age of author and age of protagonist

Null hypothesis	Sample size	Test	P-value
Proportion of authors who write about protagonists in or below their decade age range (88%) $\leq$ proportion of authors who write about protagonists who are in an older age bracket than they are (12%)	368 stories	Binomial	$<10^{-3}$

Table 17 Tests of relationship between gender of author and gender of secondary characters

Null hypothesis	Sample size	Test	P-value
Proportion of male secondary characters in stories by males (48%) $\geq$ proportion of female secondary characters in stories by males (52%)	345 characters	Binomial	$<10^{-3}$
Proportion of female secondary characters in stories by females (48%) $\geq$ proportion of male secondary characters in stories by females (52%)	198 characters	Binomial	$<10^{-3}$
Proportion of male secondary characters in stories by males (48%) $\leq$ proportion of female secondary characters in stories by females (48%)	Male: 345 Female: 198	Proportion	0.451

Table 18 Tests of relationship between country of origin of author and country of origin of secondary characters

Null hypothesis	Sample size	Test	*P*-value
Proportion of authors who write about secondary characters from their country of origin (75%) ≤ proportion of authors who write about secondary characters who do not share their country of origin (25%)	477 characters	Binomial	$<10^{-3}$
Proportion of American authors who write about secondary characters from America (84%) ≤ proportion of American authors who write about secondary characters who do not share their country of origin (16%)	316 characters	Binomial	$<10^{-3}$
Proportion of non-American authors who write about secondary characters from their country of origin (59%) ≤ proportion of non-American authors who write about secondary characters who do not share their country of origin (41%)	161 characters	Binomial	$<10^{-3}$
Proportion of American authors who write about secondary characters from their country of origin (84%) ≤ proportion of non-American authors who write about secondary characters from their country of origin (59%)	American: 316 Other: 161	Proportion	$<10^{-3}$

Table 19 Tests of relationship between race of author and race of secondary characters

Null hypothesis	Sample size	Test	*P*-value
Proportion of Caucasian secondary characters in stories by Caucasians (92%) ≤ proportion of non-Caucasian secondary characters in stories by Caucasians (8%)	388 characters	Binomial	$<10^{-3}$
Proportion of black secondary characters in stories by blacks (57%) ≤ proportion of non-black secondary characters in stories by blacks (43%)	21 characters	Binomial	$<10^{-3}$
Proportion of Asian secondary characters in stories by Asians (90%) ≤ proportion of non-Asian secondary characters in stories by Asians (10%)	40 characters	Too little data to use Binomial	NA
Proportion of Hispanic secondary characters in stories by Hispanics (83%) ≤ proportion of non-Hispanic secondary characters in stories by Hispanics (17%)	12 characters	Too little data to use Binomial	NA

Table 20 Tests of relationship between age of author and age of secondary characters

Null hypothesis	Sample size	Test	*P*-value
Proportion of authors who write about secondary characters in or below their decade age range (81%) ≤ proportion of authors who write about secondary characters who are in an older age bracket than they are (19%)	454 characters	Binomial	$<10^{-3}$

Table 21 Tests of relationship between race of author and race of characters

Null hypothesis	Sample size	Test	*P*-value
Proportion of Caucasian characters in stories by Caucasians (94%) ≤ proportion of non-Caucasian characters in stories by Caucasians (6%)	696 characters	Binomial	$<10^{-3}$
Proportion of black characters in stories by blacks (69%) ≤ proportion of non-black characters in stories by blacks (31%)	35 characters	Binomial	$<10^{-3}$
Proportion of Asian characters in stories by Asians (90%) ≤ proportion of non-Asian characters in stories by Asians (10%)	68 characters	Binomial	0.001
Proportion of Hispanic characters in stories by Hispanics (74%) ≤ proportion of non-Hispanic characters in stories by Hispanics (26%)			

(continued)

Table 21 Continued

Null hypothesis	Sample size	Test	*P*-value
Proportion of Caucasian characters in stories by Caucasians (94%) ≤ proportion of black characters in stories by blacks (69%)	Caucasian: 696 Black: 35	Proportion	$<10^{-3}$
Proportion of Caucasian characters in stories by Caucasians (94%) ≤ proportion of Asian characters in stories by Asians (90%)	Caucasian: 696 Asian: 68	Proportion	0.123
Proportion of Caucasian characters in stories by Caucasians (94%) ≤ proportion of Hispanic characters in stories by Hispanics (74%)	Caucasian: 696 Hispanic: 23	Proportion	$<10^{-3}$
Proportion of black characters in stories by blacks (69%) ≥ proportion of Asian characters in stories by Asians (90%)	Black: 35 Asian: 68	Proportion	0.008
Proportion of black characters in stories by blacks (69%) ≥ proportion of Hispanic characters in stories by Hispanics (74%)	Black: 35 Hispanic: 23	Proportion	0.443
Proportion of Asian characters in stories by Asians (90%) ≤ proportion of Hispanic characters in stories by Hispanics (74%)	Asian: 68 Hispanic: 23	Proportion	0.063

Table 22 Tests of relationship between native world region of author and world region of story setting

Null hypothesis	Sample size	Test	*P*-value
Proportion of authors who write stories set in their home region of the world (80%) ≤ proportion of authors who write stories set outside of their home region of the world (20%)	383 stories	Binomial	$<10^{-3}$
Proportion of American authors who write stories set in America (85%) ≤ proportion of American authors who write stories set outside of the US (15%)	241 stories	Binomial	$<10^{-3}$
Proportion of non-American authors who write stories set in their home region of the world (73%) ≤ proportion of non-American authors who write stories set outside of their home region of the world (17%)	142 stories	Binomial	$<10^{-3}$
Proportion of American authors who write stories set in America (85%) ≤ proportion of non-American authors who write stories set in their home region of the world (73%)	American: 241 Other: 142	Proportion	0.005