Gender associations in the twentieth-century English-language literature

Daniel Schulz[1], Štěpán Bahník[2]*

[1] Department of Psychology, University of Würzburg, Röntgenring 10, Würzburg, 97070, Germany.

[2] Faculty of Business Administration, University of Economics, Prague, náměstí Winstona Churchilla 4, Prague, 130 67, Czech Republic.

* Corresponding author, bahniks@seznam.cz

Abstract

The data required to study stereotypes held in the past are often not available. Using Google Books Ngram corpus, we explored the depiction of male and female characters in the twentieth-century English-language fiction. By analyzing adjective-noun bigrams, we examined adjectives used in association with “man”, “woman”, “boy”, and “girl”. Men were described in more positive terms than women. Girls were depicted in more positive terms than boys at the beginning of the twentieth century, but the tendency reversed in the middle of the century. Boys were described in more masculine terms than girls; however, men were described in similarly masculine adjectives as women. Despite limitations of interpretability of the results, the study presents a possible approach of exploring past stereotypes.

Keywords

gender, literature, stereotypes, English, 20th century

Acknowledgements

The study was supported by a grant of the Internal Grant Agency of the University of Economics, Prague awarded to Štěpán Bahník (IP310087). Both DS and ŠB designed the study and wrote the report. DS analyzed the data.

People have different associations with men and women, which form a basis of their gender stereotypes. Even though stereotypes differ between various gender subgroups (Fiske, Xu, Cuddy, & Glick, 1999), men tend to be generally seen as more competent and women as more warm (Kite, Deaux, & Haines, 2008). Women have also been found to be perceived generally more favorably than men (Eagly, Mladinic, & Otto, 1991). Given that gender stereotypes tend to be accurate (Swim, 1994), people may learn them through observation and induction. They are also transmitted by social learning and communicated by family, peers, and media (Kite et al., 2008).

Studies examining the depiction of men and women in media have shown that gender stereotypes are mirrored in various forms of media. A recent meta-analysis of studies of

gender depiction in advertisements showed that women are more likely to be portrayed at home, associated with domestic products, not speaking, and in a dependent role (Eisend, 2010). In children's cartoons, male characters are more likely to be portrayed as independent, assertive, athletic, and technical, while female character are more likely to be portrayed as emotional, warm, affectionate, and frail (Thompson & Zerbinos, 1995). Similarly, male characters in popular movies are more likely than female characters to have a leadership role, occupational power over others, and goals (Lauzen & Dozier, 2005). In an analysis of adjectives used to describe characters in prize-winning children's books from 1984 to 1994, Turner-Bowker (1996) found that the adjectives used for male characters were evaluated as more masculine, active, and potent even though the adjectives used to describe female characters were evaluated more positively.

Apart from the difference in depiction of male and female characters in various media, the proportion of characters of the two genders also differs. Even though the ratio of male to female characters in children's books varied over the twentieth century, the characters were more likely to be male than female for most of the century. Only in 1910s and 1990s was the proportion of male and female characters approximately equal (McCabe, Fairchild, Grauerholz, Pescosolido, & Tope, 2011). Women have also been found to be underrepresented among television characters (Elasmar, Hasegawa, & Brain, 1999), in movies (Lauzen & Dozier, 2005), in fiction books (Underwood, Bamman, & Lee, 2018), and in children's cartoons (Thompson & Zerbinos, 1995).

Recent creation of the Google Books Ngram corpus (Michel et al., 2011) has enabled the study of depiction of male and female characters in a large volume of the fiction literature. Google Books Ngram corpus contains data about word usage from more than 5 million books. The data relate to *n-gram*s, which are sequences of *n* words used in a text. Previous studies used the corpus to study, for example, the use of emotion-related words in books (Acerbi, Lampos, Garnett, & Bentley, 2013), the use of individualistic words and phrases (Twenge, Campbell, & Gentile, 2012a), or censorship of topics and regularization of irregular verbs (Michel et al., 2011). Some studies have also explored issues related to gender. Mason, Kuntz, and McGill (2015) found that women appeared in English books less often than men and that, in comparison to women, men tended to be relatively more often described as "young" than "old". Twenge, Campbell, and Gentile (2012b) found that male pronouns occurred more often in U.S. books than female pronouns and that the ratio of their use was highest during the 1950s and early 1960s, and that it has been decreasing from then on. Finally, Ye, Cai, Chen, Wan, & Qian (2018), who used a similar method as the present study, found that in English-language books, men tended to be more often described in terms associated with agreeableness, extraversion, conscientiousness, and neuroticism traits of the Big Five model and that they were described similarly often as women in terms associated with openness. However, Ye et al. (2018) did not take into account that men were generally relatively more often described with an adjective than women, which could have confounded some of their results.

In the present study, we use the Google Books Ngram corpus to explore the depiction of men and women in the English-language literature and its development in the 20th century. In particular, we focus on adjectives that are used in association with nouns "man", "woman", "boy", and "girl". The four nouns were selected to study both the effect of gender and a possible difference between description of adults and children.The data on historical development of the perception of men and women are scarce and difficult to obtain

retrospectively, so we use a simple way to study it indirectly by examining the depiction of men and women in books, which are likely to reflect contemporary views. Furthermore, the depiction of men and women in books is of interest in itself given that it may influence opinions and views of readers.

Method[1]

Bigrams

We searched for adjective-noun bigrams in the Google Books Ngram data (http://storage.googleapis.com/books/ngrams/books/datasetsv2.html). In particular, we used the English fiction subset of the dataset (version 20120701). Raw data were downloaded and extracted using a custom-written Python program (https://osf.io/egpr5/wiki/Extraction%20Program/). We used syntactic annotations (Lin et al., 2012) to select only bigrams in which adjectives (described below) were related to the nouns we studied — "man", "woman", "boy", and "girl" (as well as "person" and "child", which were however used only for the analysis of similarity of descriptions). We used only the data from years between 1900 and 2000. The precise cutoff was arbitrary, but the years preceding 1900 tended to have insufficient amount of data and after 2000 the composition of books changed (http://www.culturomics.org/Resources/faq#dataquality); the books included in the database published after 2000 no longer came predominantly from libraries.

Adjectives

We used a list of adjectives describing people's traits collected by Anderson (1968). In his study, Anderson used 18,000 trait names compiled by Allport and Olbert (1936) and reduced them to a list of 555 words by including only appropriate words used frequently for describing other people. To further reduce this list of 555 trait words, we selected 100 adjectives that were most often used in books in combination with each of the four studied nouns between years 1900 and 2008.[2] Since the adjectives did not completely overlap, we ended up with 155 adjectives. The full list of adjectives can be found in Appendix.

Three hundred thirty-nine US Amazon Mechanical Turk workers[3] rated the adjectives on a randomly chosen dimension out of masculinity, competence, desirability, and warmth. The adjectives were rated on a 7-point scale adapted from Anderson (1968) in terms of how much they are associated with femininity (0) or masculinity (6), incompetence (0) or competence (6), and coldness (0) or warmth (6), or how much they are undesirable (0) or

[1] Data, analysis code, and the program used to download and extract raw data can be found on: https://osf.io/egpr5/. The study was not pre-registered.

[2] We based the selection on the range ending with 2008 because we originally intended to include years up to 2008 in the study and we did not obtain ratings for the few adjectives that would have been selected if we based the selection on the range ending in 2000.

[3] Data from participants who failed to answer at least one of two control items instructing them to pick a specific rating (n = 64) and from participants who finished the survey in less than 200 seconds or did not answer all questions (n = 66) were excluded from analysis.

desirable (6).[4] The evaluation of the adjectives in terms of their desirability was stable between the study by Anderson (1968) and our survey (see Table 1). The stability of ratings over 49 years suggests that the adjectives kept similar connotations over the time. The adjective ratings for dimensions desirability, competence, and warmth correlated strongly between each other (see Table 1), so we computed their averages and used the composite score in analysis (the dimension called "positivity"). A positive correlation of traits in terms of their association with competence and warmth has been observed before and it has been shown to be due to the shared association with positive valence (Suitner & Maas, 2008).

Variable	*M*	*SD*	1	2	3	4	5
1. Anderson (1968)	3.50	1.41					
2. Positivity	3.40	1.37	.95**				
3. Masculinity	2.97	0.77	-.28**	-.25**			
4. Desirability	3.49	1.78	.95**	.99**	-.22**		
5. Competence	3.41	1.30	.89**	.95**	-.10	.95**	
6. Warmth	3.30	1.20	.89**	.92**	-.42**	.89**	.75**

Table 1. Correlations between the dimensions used for adjective ratings. The table shows correlations between adjective ratings on four dimensions obtained in the present study (3-6), likeability ratings from Anderson (1968; 1), and the composite positivity score (2) computed from ratings on dimensions 4-6. *Note:* ** indicates $p < .01$.

Composite scores

The extraction yielded in total 5,795,481 syntactic bigrams over the four nouns combined. To study the description of men, women, boys, and girls in terms of positivity and masculinity, we used composite scores computed for each combination of a noun, dimension, and year (e.g., the positivity of adjectives used for depiction of "man" in the year 1950). To obtain the composite scores, we computed a weighted average of adjective ratings. That is, we multiplied average ratings on the dimension of the 100 adjectives most frequently associated with the noun with the proportion of bigrams from the given year that contained the adjective-noun bigram. For the year *y*, dimension *d*, and noun *n*, the composite score *CS* was thus computed as:

[4] We actually used a scale from 1 to 7, but we subtracted 1 from the ratings to make the results comparable to Anderson (1968) who used the scale from "least favorable or desirable" (0) to "most favorable or desirable" (6).

$$CS_{y,d,n} = \frac{\sum_{a \in A_n} f_{a,y,d,n} \times AR_{a,d}}{\sum_{a \in A_n} f_{a,y,d,n}},$$

where *f* is frequency, *a* is an adjective, *A* is the set of the most frequent adjectives, and *AR* is the average rating of an adjective. We used these composite scores in subsequent analyses.

Results

Throughout the studied period of time there were more bigrams describing men than women and more bigrams describing girls than boys (see Figure 1).

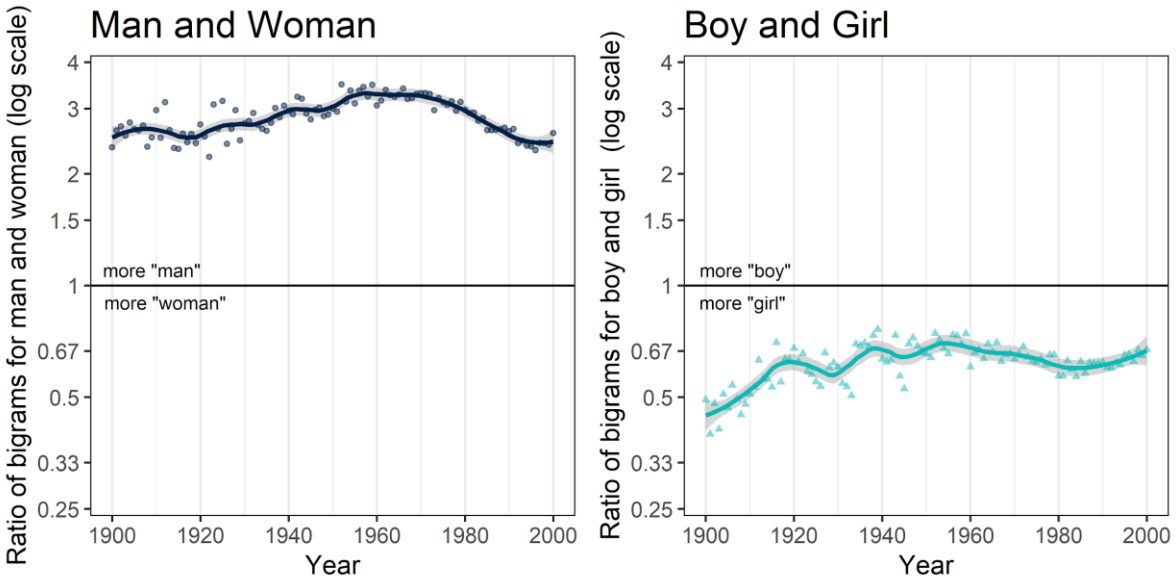

Figure 1. Ratio of bigrams describing men and women (left) and boys and girls (right). The lines show results of loess regression fitted for the relative frequencies of bigrams. The shaded regions represent 95% confidence intervals for the regression line. Note that the ordinate uses a logarithmic scale.

To model the trends in the description of genders in the literature, we used a segmented regression for each noun (or difference between two nouns) and dimension (Muggeo, 2008). The number of breakpoints in each model was selected using the Bayesian information criterion (BIC). Afterward, a segmented regression model was fitted with the number of breakpoints determined by the BIC. The resulting models for the difference between the genders are displayed in Figure 2 and models for the two genders separately are displayed in Figure 3.

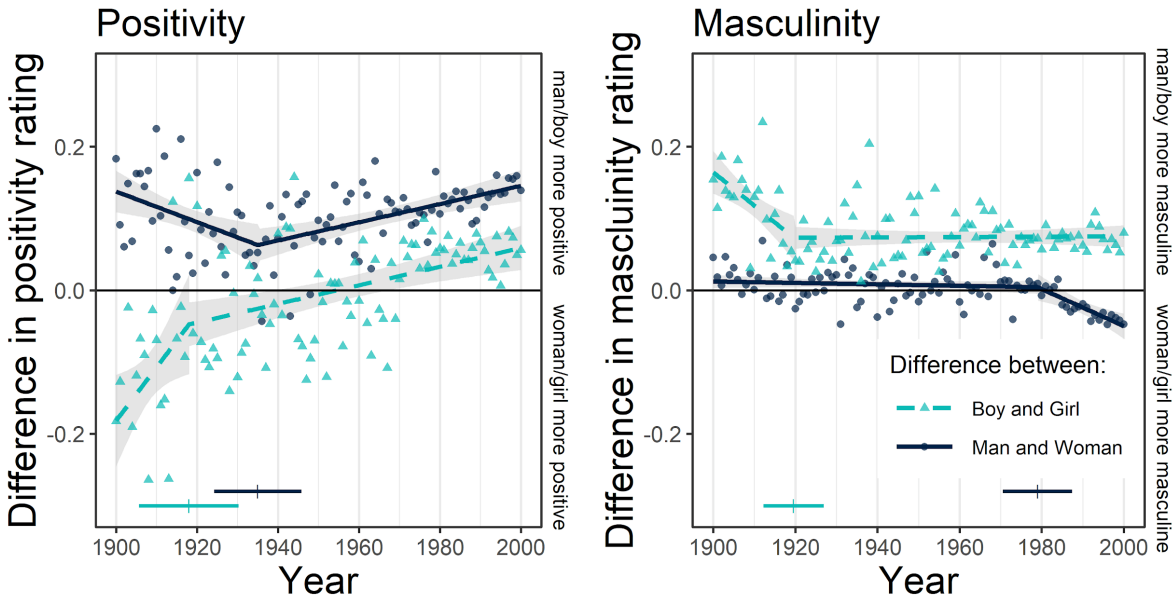

Figure 2. The difference in positivity and masculinity associated with adjectives used for description of men and women, and of boys and girls. The lines represent results of segmented regressions and shaded regions show 95% confidence intervals for the regression lines. Lines at the bottom of the graph show 95% confidence intervals for the breakpoints.

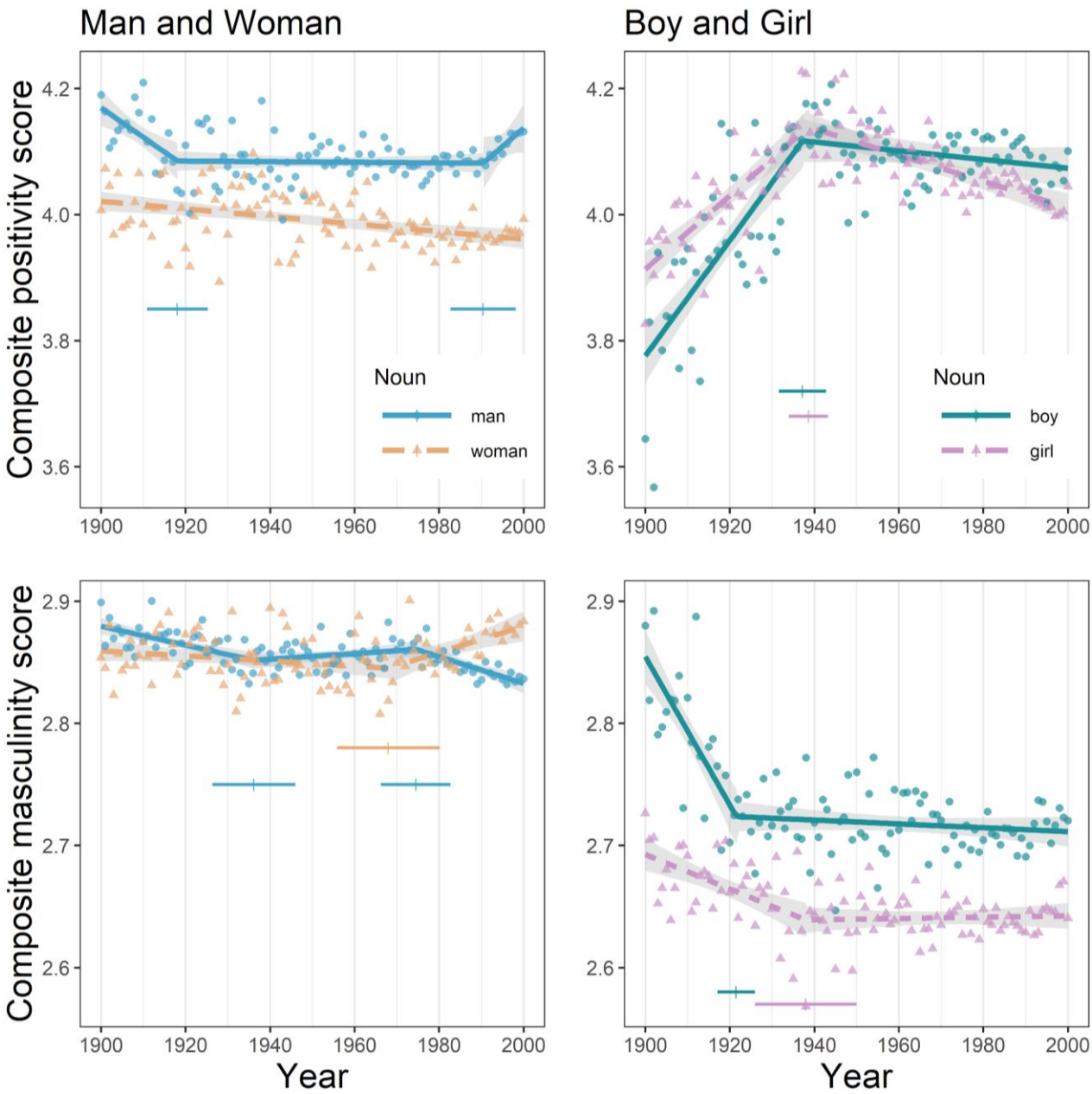

Figure 3. Positivity and masculinity associated with adjectives used for description of men, women, boys, and girls. The lines represent results of segmented regressions and shaded regions show 95% confidence intervals for the regression lines. Horizontal lines below the data points show 95% confidence intervals for the breakpoints.

Both genders were generally described in positive terms; that is, well above the midpoint of the scale (3). Men were generally depicted in more positive terms than women, $t(198) = -19.09$, $p < .001$, $b = -0.10$, 95% CI = [-0.11, -0.09]. Men were more likely to be depicted as "honest" and less likely to be described as "foolish", "unhappy", "jealous", "vulgar", and "silly", which most influenced the higher average positivity of description of men. On the other hand, men were more likely to be described as "lazy" and "mean" and less likely to be described as "charming", which had the opposite, but overall weaker, influence (see Table 2). The difference in positivity decreased until about 1937 and increased back afterward. While men were depicted in similarly positive terms throughout most of the 20th century (1913-1989)

with an exception of the decrease at the beginning and increase at the end of the century, positivity of description of women was decreasing throughout the century.

Rank	Positivity		Masculinity	
	man-woman	**boy-girl**	**man-woman**	**boy-girl**
1st	honest [M] (-0.0246)	good [B] (-0.1475)	charming [W] (+0.017)	good [B] (+0.0278)
2nd	foolish [W] (-0.0239)	nice [G] (+0.0199)	wise [M] (-0.0100)	charming [G] (+0.0262)
3rd	unhappy [W] (-0.0174)	bright [B] (-0.0199)	vulgar [W] (+0.0091)	innocent [G] (-0.0195)
4th	charming [W] (+0.0168)	charming [G] (+0.0163)	innocent [M] (+0.0087)	nice [G] (-0.0101)
5th	jealous [W] (-0.0148)	foolish [B] (+0.0161)	fashionable [W] (-0.0077)	sensitive [B] (+0.0089)
6th	vulgar [W] (-0.0131)	unhappy [G] (-0.0155)	independent [W] (+0.0073)	tough [B] (-0.0056)
7th	good [M] (+0.0127)	dull [B] (+0.0124)	honorable [M] (-0.0057)	bright [B] (-0.0048)
8th	silly [W] (-0.0110)	sensible [G] (+0.0103)	able [M] (-0.0055)	proud [G] (+0.0047)
9th	lazy [M] (+0.0099)	inexperienced [G] (-0.0102)	warm [W] (-0.0052)	reckless [B] (-0.0038)
10th	mean [M] (+0.0098)	cruel [B] (+0.0100)	foolish [W] (+0.0046)	cruel [B] (-0.0035)
Absolute difference	0.1032	-0.0165	0.0014	0.0831

Table 2. The adjectives that had most influence on the overall differences between the two genders. The table displays ten adjectives with highest influence scores for all combinations of the two pairs of nouns and the two dimensions. The influence scores for each adjective were computed by subtracting differences in the composite scores with and without the adjective. That is, the influence score shows how much would the difference in composite scores for a given dimension (shown in the last row) change if a given adjective was not included in the composite scores. The letters in brackets show which of the two nouns was relatively more often associated with the adjective.

Both genders were described in slightly feminine terms; that is, on average below the midpoint of the scale (3). There was generally little difference in masculinity of adjectives used for description of men and women, $t(198) = -0.64$, $p = .53$, $b = -0.00$, 95% CI = [-0.01, 0.00], with a possible exception of the end of the century when men tended to be described

in somewhat less masculine adjectives than women. Men were more likely to be described as “wise”, “honorable”, and “able”, all relatively masculine words, and less likely to be described as “fashionable” and “warm”, both relatively feminine words. On the other hand, women were more likely to be described as “charming”, “vulgar”, “independent”, and “foolish”, all rated as relatively masculine traits. Men were described in decreasingly masculine terms throughout the beginning (1900-1937) and end (1975-2000) of the 20th century with little change between these two periods. Masculinity of adjectives used to describe women did not change throughout the first half of the century (1900-1968), but it started to slightly increase afterward (1968-2000).

Positivity of adjectives showed the same general trends for both boys and girls. While they were described in increasingly more positive terms until around 1938, the positivity was slightly decreasing from then on. The increase in positivity at the beginning of the century was larger for boys and the subsequent decrease was smaller for boys, resulting in a reversal of the relative positivity of adjectives used in relation to boys and girls. While girls were described more positively at the beginning of the century, the difference was changing in favor of boys throughout the studied period and boys started to be described in more positive terms since around 1960. The relative change of the difference was especially prominent at the beginning of the 20th century (1900-1918). In aggregate, boys were described in similarly positive terms as girls, $t(198) = 1.41$, $p = .16$, $b = 0.02$, 95% CI = [-0.01, 0.04]. Boys were more likely to be described as “good”, which had the most influence on the difference in positivity of descriptions of boys and girls given the overall frequency of the corresponding bigrams (31.7% for boys and 22.5% for girls). To a smaller degree, boys were depicted more positively also because they were more likely to be described as “bright” and less likely to be described as “unhappy” and “inexperienced”. On the other hand, boys were more likely to be described by the negative adjectives “foolish”, “dull”, and “cruel” and less likely to be described by the positive adjectives “nice”, “charming”, and “sensible” which decreased the difference in positivity of description of the two genders (Table 2).

Children were depicted in more feminine terms than adults. Boys were described generally in more masculine adjectives than girls, $t(198) = -19.25$, $p < .001$, $b = -0.08$, 95% CI = [-0.09, -0.07]. Boys were more likely to be depicted as “tough”, “bright”, “reckless”, and “cruel”, all rated as relatively masculine traits, and less likely to be described as “innocent” and “nice”, both rated as feminine. On the other hand, girls were more likely to be described as “charming”, rated as relatively masculine, and less likely to be described as “good” and “sensitive”, both rated as feminine traits. The difference in masculinity of descriptions decreased at the beginning of the 20th century (1900-1920) and then stayed at the same level over the studied period of time. Masculinity of adjectives used for description of boys and girls displayed a similar development in time. While it was decreasing until about 1922 for boys and 1938 for girls, it did not change appreciably until the end of the century.

Finally, we examined the similarity of adjectives used to describe characters of the two genders. Namely, we computed Euclidean distance of the proportions of adjectives used to describe pairs of nouns (see Figure 4). The analysis showed that men and women were generally described in more similar terms than boys and girls, $t(198) = 16.60$, $p < .001$, $b = 0.040$, 95% CI = [0.035, 0.045]. While the similarity of description of girls and boys did not change appreciably during the century, men and women were described increasingly more similarly until about 1977 when the trend reversed. Women were described generally more similarly to “person” than men, $t(198) = 2.56$, $p = .01$, $b = 0.005$, 95% CI = [0.001, 0.009], but

the difference actually occurred only in the second half of the century. Moreover, the terms used to describe “child” were more similar to the terms used to describe girls than to the terms used to describe boys, $t(198) = 14.50$, $p < .001$, $b = 0.053$, 95% CI = [0.046, 0.060]. Women and girls tended to be described more similarly than men and boys, $t(198) = 16.25$, $p < .001$, $b = 0.039$, 95% CI = [0.034, 0.044], but the difference seemed to disappear at the end of the century.

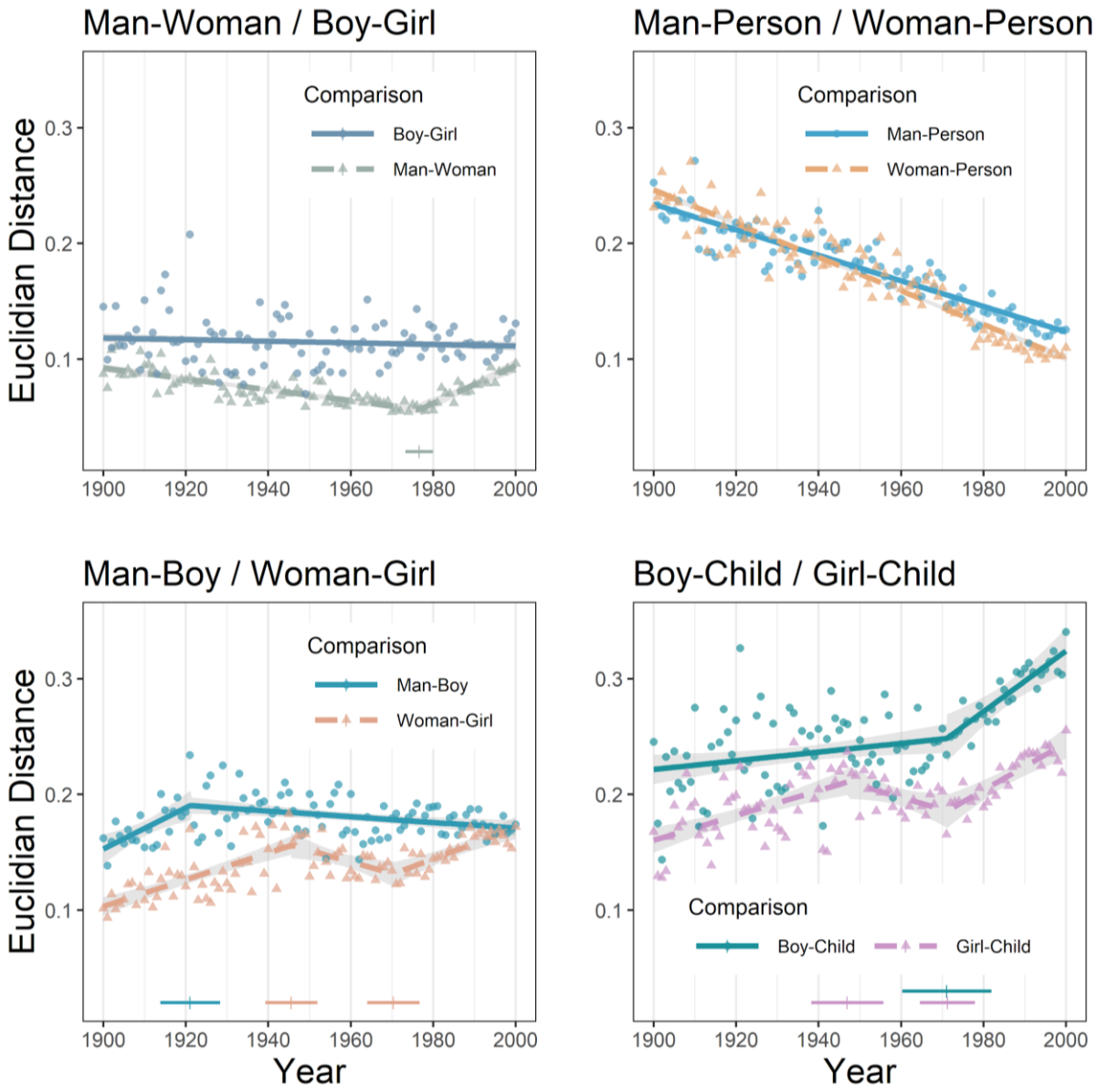

Figure 4. Euclidean distances of proportions of adjectives used for description of pairs of nouns. The lines represent results of segmented regressions and shaded regions show 95% confidence intervals for the regression lines. Horizontal lines below the data points show 95% confidence intervals for the breakpoints.

Discussion

Similarly as previous studies examining children’s books and other media, we found that men were more often portrayed in the English fiction than women. As McCabe et al. (2011)

observed for children's books and Twenge et al. (2012b) observed for the relative use of gendered pronouns in U.S. books, the difference was lower at the beginning and end of the 20th century, but unlike in children's books the frequency of portrayal of men in general fiction was always at least twice as high as that of women. The same pattern of results for characters depicted in books has been recently identified by Underwood et al. (2018) using a different method for identifying characters and a different source of data, supporting validity of the approach used in the present study. Underwood et al. (2018) argue that the observed pattern may be caused by the proportion of women authors which was at its lowest around 1970. Women tend to include women characters relatively more often, so the nadir of women authors around 1970 could explain the higher male-to-female ratio of characters in books. Here, we show that the overrepresentation of male characters did not hold for children, where "boy" was less likely to occur in a bigram than "girl". Yet, given that the adult nouns were much more frequent than the children nouns, male characters were represented in the books more often than female characters overall. Future studies may explore whether the difference between adult and children characters is present also in other media.

While some previous research found that women are evaluated more positively than men (Eagly et al., 1991), we observed the opposite result in characterization of men and women in the literature. Men were generally described in more positive terms. The difference also did not decrease throughout the century. Men were still described more positively than women even at the end of the century despite the, in many aspects, increased status of women in the society. As a speculation, it is possible that men tend to describe men relatively more positively and the higher number of male authors could cause the relatively more positive description of male characters. It is also possible that male characters are more likely to be given central roles and main characters are more likely to be described positively than supporting characters. The pattern of results differed for girls and boys. While girls were described in more positive terms at the beginning of the century, the difference gradually reversed throughout the century.

The lack of difference in the average masculinity of adjectives describing men and women seems surprising. One possible reason for this finding may be that feminine characteristics of men and masculine characteristics of women may be more likely to be mentioned because they are more likely to be informative. Describing a feature of a person that is expected is not as informative as describing a feature that goes against expectations and which therefore distinguishes the person better. For example, "male nurse" is used more often than "female nurse" even though nurses are more likely to be female than male. However, the expected difference in masculinity was observed for boys and girls, which suggests that the association of the adjectives with masculinity tracks a real feature. It is also noteworthy that adjectives used for description of children were generally less associated with masculinity than adjectives used for description of adults. It has been found that children are rated to have less agency (Gray, Gray, & Wegner, 2007), which could have been tracked by the masculinity dimension (Spence & Helmreich, 1980).

The interpretation of trends in depiction of the two genders is not only complicated by the effect of informativeness of adjectives, which may influence the relative frequency of their use for describing the two genders, but also by difference in meaning of the adjectives when they are used to describe men and women. For example, the height necessary for people to be referred to as "tall man" and "tall woman" might be quite different. It is similarly possible that "angry woman" and "angry man" refer to different levels of anger. This could explain the

difference between self-reported traits in actual people and traits by which men and women are described in the literature (Ye et al., 2018).

Underwood et al. (2018) showed that over the 20th century gender of fiction characters became harder to classify based on their description and associated words. This could suggest that characters in books became less stereotypically associated with their own genders. We did not see a similar convergence in masculinity of adjectives used to depict men and women in our data because there was no difference in their masculinity from the beginning of the 20th century. Only in children nouns did we observe a closing gap between genders, but it could be seen only at the beginning of the 20th century until about 1920. However, the analysis of similarity of description of men and women showed increasing similarity for most of the 20th century, which could partly explain the increasing difficulty of classification of characters based on gender observed by Underwood et al. However, their study did not observe a similar reversal of the trend in the last quarter of the century that was found in our data. The similarity of description of boys and girls also did not change appreciably over the century.

Even though it is possible to study the use of words and combinations of words in the literature using the Google Books Ngram corpus, it is often difficult to interpret what the results mean. For example, the observed differences in frequencies of bigrams using “man” and “woman” or “boy” and “girl” suggest overrepresentation of male adult characters and female children characters in the literature; however, it is not clear whether the difference is caused by the number of characters of both genders, their prominency in books, or both. Examining only the development of frequencies of bigrams without their historical context also means that it is hard to interpret what was the cause of changes in representation of the genders. The main results are therefore mostly descriptions of trends and their explanations have to be studied differently.

The lack of specific tested effects of historical events means that we examined only long-term trends rather than short-term fluctuations. It is possible that some short periods of time (e.g., the world wars) significantly influenced the depiction of characters in the literature, but the effect was obscured by the analysis we used. It is therefore not necessarily true that the results we describe apply in the same manner to the whole studied period. On the other hand, it is possible that most short-term events are not likely to influence the depiction of characters immediately and they would not show in the results as a short-term effect anyway. Bentley, Acerbi, Ormerod, Lampos (2014), for example, argue that the condition of economy influences language used in the literature most with a ten-year delay.

The composite scores of positivity and masculinity were computed from ratings of adjectives from the perspective of present people. The comparison of ratings of desirability from almost half a century ago (Anderson, 1968) with ratings obtained in the present study showed a near-perfect correlation, suggesting that the word associations are largely stable; yet, insofar that the meaning of some of the words changed (Pettit, 2016), the composite scores for farther past could be biased. Using a list of words from Anderson (1968) also means that some of the deprecated older words as well as some novel words could not have been included even if they were frequently used in a certain time period.

The corpus does not contain all published fiction books and it is not known to what degree are the included books representative of all the books (Koplenig, 2017). Furthermore, the

fiction corpus does not contain only works of fiction, but also some works associated with fiction such as commentaries (http://www.culturomics.org/Resources/faq#dataquality). However, a small proportion of texts that do not fall in the fiction category is unlikely to largely influence our results. The fiction corpus has been shown to be less influenced by publishing trends than the general corpus, which is significantly influenced by changing proportion of scientific texts (Pechenick, Danforth, & Dodds, 2015). While the changing proportion of fiction books on all books may not be an issue for the present study, the changing composition of different genres of fiction may be behind some of the results. Some genres may be more likely to include male or female characters and portray them differently than other genres. The observed trends might then be caused by the changing composition of the corpus in terms of literary genres rather than by a change of portrayal of characters within any given genre. Furthermore, the results do not necessarily correspond to the change of depiction of characters that readers encountered during a given time period. Each book is in the corpus only once, independent of its popularity (Pechenick et al., 2015). More popular books that had larger influence on gender perception could have differed in their portrayal of characters from less popular books which had lesser influence on the culture at a given time period. The characters that readers actually read about would then differ from the average characters depicted in all books.

Even though the interpretation of the results is associated with certain limitations, the present research shows a possible approach for studying depiction of the two genders in the literature. Unlike most of the previous studies using the Google Books Ngram corpus, our study did not pick specific ngrams for analysis in an ad hoc fashion, and it is therefore not influenced by bias in selection of particular words or phrases as some previous studies could have been. Nevertheless, the limitations mean that the results need to be replicated using different data or methods that could corroborate the current results.

References

Acerbi, A., Lampos, V., Garnett, P., & Bentley, R. A. (2013). The Expression of Emotions in 20th Century Books. *PLOS ONE*, *8*(3), e59030.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*(1), i-171.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*(3), 272–279.

Bentley, R. A., Acerbi, A., Ormerod, P., & Lampos, V. (2014). Books average previous decade of economic misery. *PlOS ONE*, *9*(1), e83147.

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men?. *Psychology of Women Quarterly*, *15*(2), 203-216.

Eisend, M. (2010). A meta-analysis of gender roles in advertising. *Journal of the Academy of Marketing Science*, *38*(4), 418-440.

Elasmar, M., Hasegawa, K., & Brain, M. (1999). The portrayal of women in US prime time television. *Journal of Broadcasting & Electronic Media*, *43*(1), 20-34.

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis)respecting versus (dis)liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473-489.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619.

Kite, M. E., Deaux, K., & Haines, E. L. (2008). Gender stereotypes. In F. L. Denmark & M. A. Paludi (Eds.), *Psychology of women: A handbook of issues and theories* (2nd ed., pp. 205–236). Westport, CT: Praeger.

Koplenig, A. (2017). The impact of lacking metadata for the measurement of cultural and linguistic change using the Google Ngram data sets—Reconstructing the composition of the German corpus in times of WWII. *Digital Scholarship in the Humanities*, *32*(1), 169-188.

Lauzen, M. M., & Dozier, D. M. (2005). Maintaining the double standard: Portrayals of age and gender in popular films. *Sex Roles*, *52*(7), 437-446.

Lin, Y., Michel, J. B., Aiden, E. L., Orwant, J., Brockman, W., & Petrov, S. (2012, July). Syntactic annotations for the google books ngram corpus. In *Proceedings of the ACL 2012 system demonstrations* (pp. 169-174). Association for Computational Linguistics.

Mason, S. E., Kuntz, C. V., & McGill, C. M. (2015). Oldsters and Ngrams: Age stereotypes across time. *Psychological Reports*, *116*(1), 324-329.

McCabe, J., Fairchild, E., Grauerholz, L., Pescosolido, B. A., & Tope, D. (2011). Gender in twentieth-century children's books: Patterns of disparity in titles and central characters. *Gender & Society*, *25*(2), 197-226.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., … Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182.

Muggeo, V. M. R. (2008). segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News*, *8*(1), 20–25.

Pechenick, E. A., Danforth, C. M., & Dodds, P. S. (2015). Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PLOS ONE*, *10*(10), e0137041.

Pettit, M. (2016). Historical time in the age of big data: Cultural psychology, historical change, and the Google Books Ngram Viewer. *History of Psychology*, *19*(2), 141-153.

Spence, J. T., & Helmreich, R. L. (1980). Masculine instrumentality and feminine expressiveness: Their relationships with sex role attitudes and behaviors. *Psychology of Women Quarterly*, *5*(2), 147-163.

Suitner, C., & Maass, A. (2008). The role of valence in the perception of agency and communion. *European Journal of Social Psychology*, *38*(7), 1073-1082.

Swim, J. K. (1994). Perceived Versus Meta-Analytic Effect Sizes: An Assessment of the Accuracy of Gender Stereotypes. *Journal of Personality and Social Psychology*, *66*(1), 21-36.

Thompson, T. L., & Zerbinos, E. (1995). Gender roles in animated cartoons: Has the picture changed in 20 years? *Sex Roles*, *32*(9-10), 651-673.

Turner-Bowker, D. M. (1996). Gender stereotyped descriptors in children's picture books: Does “curious Jane” exist in the literature? *Sex Roles*, *35*(7), 461-488.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012a). Increases in individualistic words and phrases in American books, 1960–2008. *PLOS ONE*, *7*(7), e40181.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012b). Male and female pronoun use in US books reflects women’s status, 1900–2008. *Sex Roles*, *67*(9-10), 488-493.

Underwood, T., Bamman, D., & Lee, S. (2018). The Transformation of Gender in English-Language Fiction. *Journal of Cultural Analytics*. doi: 10.22148/16.019

Ye, S., Cai, S., Chen, C., Wan, Q., & Qian, X. (2018). How have males and females been described over the past two centuries? An analysis of Big-Five personality-related adjectives in the Google English Books. *Journal of Research in Personality*, *76*, 6-16.

Appendix

adjective	competence	desirable	masculinity	warmth	positivity	man	woman	boy	girl
able	5.10	5.13	3.70	4.00	4.74	*0.99%*	*0.38%*	*0.16%*	*0.13%*
active	4.16	4.64	3.78	3.97	4.26	*0.58%*	*0.41%*	*0.41%*	*0.14%*
admirable	4.54	5.07	3.44	4.54	4.72	0.17%	*0.49%*	0.04%	0.08%
adventurous	4.16	4.58	4.17	4.15	4.30	0.05%	0.03%	*0.14%*	0.05%
agreeable	3.91	4.56	2.31	4.63	4.37	*0.36%*	*0.41%*	0.05%	0.08%
ambitious	4.84	4.90	4.04	3.58	4.44	*0.59%*	*0.45%*	*0.39%*	*0.26%*
amiable	3.96	4.63	2.45	4.37	4.32	*0.47%*	*0.55%*	0.12%	*0.37%*
angry	1.71	0.73	4.22	1.36	1.27	*0.72%*	*0.80%*	*0.39%*	*0.25%*
anxious	2.18	1.12	2.44	2.35	1.88	*0.23%*	*0.29%*	*0.25%*	*0.21%*
average	2.99	2.64	3.12	3.25	2.96	*1.48%*	*0.78%*	*0.74%*	*0.43%*
bashful	2.61	2.17	1.92	3.08	2.62	0.05%	0.01%	*0.24%*	0.09%
bold	3.89	4.25	4.35	3.39	3.84	*0.53%*	*0.28%*	*0.41%*	*0.34%*
bright	4.92	5.25	2.97	4.29	4.82	*0.35%*	*0.65%*	*4.10%*	*1.73%*
brilliant	5.38	5.38	3.39	3.91	4.89	*0.60%*	*0.69%*	*0.59%*	*0.39%*
calm	4.13	5.03	2.59	3.93	4.36	0.21%	*0.26%*	0.06%	*0.14%*
capable	5.29	5.32	3.54	3.87	4.83	*0.89%*	*1.21%*	*0.16%*	*0.37%*
careful	4.54	4.83	2.20	3.55	4.31	*0.37%*	0.14%	0.12%	0.10%
careless	1.04	0.74	3.71	1.85	1.21	0.15%	0.10%	*0.31%*	*0.16%*
cautious	3.99	3.99	2.40	2.83	3.60	*0.33%*	0.07%	0.04%	0.02%
charming	4.10	4.85	3.48	4.79	4.58	*0.85%*	*3.29%*	*1.22%*	*3.97%*
cheerful	3.87	4.99	2.27	5.01	4.62	*0.32%*	*0.40%*	*0.26%*	*0.27%*
clean	4.16	5.19	1.89	3.76	4.37	*0.26%*	*0.32%*	*0.54%*	*0.32%*

clever	4.66	5.06	3.15	3.49	4.40	*2.30%*	*3.03%*	*2.43%*	*2.41%*
clumsy	1.51	1.04	2.88	2.57	1.71	0.09%	0.07%	*0.34%*	*0.17%*
cold	1.85	0.76	3.54	0.53	1.05	*0.37%*	*0.56%*	0.10%	*0.20%*
competent	5.65	5.39	3.71	3.99	5.01	*0.22%*	*0.30%*	0.02%	0.09%
confident	4.68	5.02	3.95	4.04	4.58	0.19%	0.19%	0.08%	0.09%
conscientious	4.68	5.06	2.46	4.29	4.68	*0.27%*	0.15%	0.08%	0.11%
courageous	4.39	5.07	4.31	4.34	4.60	*0.25%*	*0.37%*	0.11%	*0.16%*
cruel	1.42	0.39	3.92	0.68	0.83	*0.48%*	*0.48%*	*0.49%*	*0.19%*
cunning	3.73	3.09	3.48	2.21	3.01	*0.25%*	*0.21%*	0.10%	0.06%
curious	4.33	4.82	2.95	3.97	4.37	*0.22%*	*0.27%*	*0.26%*	*0.20%*
decent	4.29	5.09	2.58	4.24	4.54	*1.56%*	*1.92%*	*0.64%*	*1.52%*
dignified	4.37	4.69	3.18	3.46	4.17	*0.23%*	*0.22%*	0.02%	0.04%
dull	2.01	0.82	3.27	1.99	1.61	*0.30%*	*0.32%*	*0.75%*	*0.26%*
eager	3.70	4.26	3.07	4.08	4.01	*0.23%*	*0.28%*	*0.71%*	*0.46%*
earnest	4.11	4.67	3.22	4.37	4.38	*0.26%*	0.16%	*0.15%*	0.11%
eccentric	3.06	3.19	3.14	3.27	3.17	0.18%	0.18%	0.03%	0.03%
educated	5.16	5.36	3.29	3.73	4.75	*1.03%*	*0.47%*	0.12%	*0.30%*
efficient	5.24	5.20	3.22	3.48	4.64	0.12%	*0.23%*	0.02%	0.07%
energetic	4.15	4.69	3.34	4.28	4.37	*0.31%*	*0.40%*	0.11%	0.09%
excited	3.53	4.19	2.82	4.21	3.98	0.12%	0.17%	*0.34%*	*0.21%*
experienced	5.25	4.98	3.87	3.67	4.63	*0.31%*	*0.42%*	0.02%	0.06%
fashionable	3.48	3.79	1.00	3.52	3.60	0.09%	*0.39%*	0.00%	0.08%
foolish	1.03	0.66	3.37	1.99	1.23	*0.69%*	*1.52%*	*2.23%*	*1.67%*
frank	3.76	3.95	4.06	2.34	3.35	0.08%	0.07%	*0.18%*	*0.13%*
friendly	4.11	5.29	2.17	5.28	4.89	*0.35%*	*0.34%*	*0.27%*	*0.25%*
frivolous	1.71	1.38	2.50	2.43	1.84	0.05%	*0.23%*	0.02%	*0.20%*
generous	4.04	5.06	2.41	5.16	4.75	*0.68%*	*0.45%*	*0.40%*	*0.32%*
gentle	3.67	4.79	1.32	4.95	4.47	*0.93%*	*1.06%*	*0.88%*	*1.00%*
good	4.61	5.24	2.71	4.98	4.94	*18.98%*	*18.03%*	*30.30%*	*22.13%*
gracious	4.11	5.09	1.77	4.93	4.71	0.05%	*0.27%*	0.12%	0.07%
greedy	1.71	0.62	3.81	1.06	1.13	0.21%	*0.20%*	*0.33%*	0.11%
happy	3.97	5.22	2.68	4.95	4.71	*1.92%*	*1.50%*	*1.18%*	*1.23%*
headstrong	3.37	3.11	4.00	2.77	3.08	0.06%	0.12%	*0.25%*	*0.35%*
heartless	1.59	0.56	3.85	0.65	0.93	0.08%	*0.21%*	0.09%	0.11%
helpless	1.28	0.89	1.72	2.42	1.53	*0.33%*	*0.65%*	*0.33%*	*0.66%*
honest	4.46	5.36	2.79	4.77	4.86	*6.05%*	*2.96%*	*0.82%*	*1.14%*
honorable	4.57	5.17	3.64	4.56	4.77	*0.68%*	0.12%	0.04%	0.03%
humble	4.00	4.79	2.57	4.68	4.49	*0.31%*	*0.20%*	0.10%	*0.17%*
imaginative	4.73	5.16	2.91	4.11	4.67	0.17%	0.09%	*0.32%*	0.12%
immature	0.89	0.79	3.86	2.08	1.25	0.02%	0.02%	0.09%	0.12%
impressionable	2.61	2.78	2.21	3.61	3.00	0.05%	0.03%	*0.15%*	*0.18%*
impulsive	1.72	1.74	3.20	2.49	1.98	0.08%	0.14%	0.13%	*0.26%*
independent	4.59	5.07	3.80	3.38	4.35	0.19%	*0.72%*	0.06%	*0.18%*
inexperienced	1.13	1.43	2.33	2.59	1.72	0.09%	0.13%	*0.30%*	*0.71%*
innocent	3.10	3.67	1.66	4.38	3.72	*1.55%*	*0.89%*	*1.46%*	*3.35%*
inquisitive	4.70	4.70	2.81	3.82	4.41	0.07%	0.11%	*0.15%*	0.06%
insolent	1.52	1.00	3.29	1.87	1.46	0.05%	0.06%	0.12%	0.06%

intellectual	5.19	5.11	3.54	3.45	4.58	0.16%	*0.31%*	0.04%	0.10%
intelligent	5.29	5.55	3.14	3.78	4.87	*1.74%*	*2.29%*	*1.22%*	*1.36%*
interesting	4.29	5.21	3.11	4.14	4.55	*0.57%*	*0.59%*	*0.19%*	*0.35%*
irresponsible	0.77	0.54	3.71	1.83	1.05	0.04%	0.04%	*0.16%*	0.05%
jealous	1.75	0.64	2.28	1.60	1.33	*0.29%*	*0.82%*	0.10%	0.13%
kind	3.90	5.42	1.97	5.28	4.87	*0.25%*	*0.24%*	*0.14%*	0.12%
kindly	3.94	5.15	2.05	5.00	4.70	*0.33%*	*0.24%*	0.02%	0.04%
lazy	0.89	0.63	3.76	1.94	1.15	*0.32%*	0.16%	*0.56%*	*0.29%*
literary	4.44	4.72	2.51	3.33	4.16	*0.93%*	0.13%	0.02%	0.02%
lively	3.86	4.76	2.90	4.55	4.39	0.13%	*0.32%*	*0.48%*	*0.68%*
lonely	2.27	1.09	2.74	1.77	1.71	*1.24%*	*1.54%*	*1.30%*	*0.98%*
loyal	4.32	5.27	2.76	4.79	4.79	0.21%	0.19%	0.08%	*0.14%*
lucky	3.37	4.67	2.94	3.91	3.98	*1.08%*	*0.50%*	*0.65%*	*1.33%*
mature	4.89	5.08	2.62	3.81	4.59	*0.36%*	*1.25%*	0.04%	0.08%
mean	1.53	0.46	3.77	0.86	0.95	*0.29%*	0.16%	*0.17%*	0.06%
melancholy	2.20	1.48	2.52	2.15	1.94	*0.23%*	0.07%	*0.18%*	0.08%
modern	3.87	4.08	2.81	3.17	3.71	*1.12%*	*1.30%*	*0.14%*	*1.09%*
modest	3.86	4.57	2.13	4.24	4.22	*0.51%*	*0.50%*	*0.18%*	*0.61%*
moral	4.33	5.35	2.58	4.51	4.73	*0.27%*	0.11%	0.06%	0.02%
naive	1.49	1.40	2.06	2.81	1.90	0.06%	0.08%	*0.15%*	*0.35%*
neat	4.27	4.79	1.90	3.60	4.22	*0.28%*	*0.40%*	0.10%	*0.26%*
nervous	2.00	1.21	2.35	2.28	1.83	*0.57%*	*0.67%*	*0.43%*	*0.32%*
nice	3.95	5.21	2.03	5.00	4.72	*2.99%*	*3.24%*	*11.23%*	*13.57%*
noisy	2.04	1.25	3.62	2.32	1.87	0.06%	0.07%	*0.17%*	0.07%
normal	3.63	3.82	3.07	3.63	3.69	*0.67%*	*0.75%*	*0.95%*	*0.58%*
obedient	3.43	3.60	2.04	3.67	3.57	0.02%	0.04%	*0.29%*	*0.22%*
obstinate	2.29	1.61	3.49	2.13	2.01	*0.23%*	*0.22%*	*0.17%*	*0.13%*
ordinary	3.08	2.61	3.07	3.14	2.94	*2.32%*	*1.74%*	*1.08%*	*1.37%*
patient	4.61	5.14	1.87	4.39	4.71	*0.39%*	0.17%	0.04%	0.06%
pleasant	4.00	5.21	2.12	4.95	4.72	*0.69%*	*0.86%*	*0.31%*	*0.49%*
polite	4.19	5.23	1.95	4.56	4.66	*0.26%*	0.06%	*0.34%*	0.09%
popular	3.77	3.83	2.70	4.07	3.89	*0.44%*	0.12%	*0.32%*	*0.51%*
practical	4.76	4.73	3.45	3.38	4.29	*1.05%*	*0.66%*	0.06%	*0.17%*
proud	3.73	3.92	4.01	3.80	3.82	*1.26%*	*1.54%*	*0.62%*	*0.90%*
prudent	3.73	3.94	2.55	2.80	3.49	*0.42%*	*0.22%*	0.01%	0.06%
quick	4.25	4.18	3.59	3.15	3.86	0.18%	0.15%	*0.37%*	*0.18%*
quiet	3.15	3.17	2.23	2.76	3.03	*1.71%*	*1.26%*	*1.70%*	*1.69%*
rash	1.46	1.04	3.88	1.91	1.47	0.11%	0.02%	*0.14%*	0.05%
rational	4.87	5.13	3.62	3.61	4.54	*0.25%*	0.10%	0.00%	0.01%
reasonable	4.73	5.21	2.91	4.17	4.70	*0.76%*	*0.35%*	0.04%	0.07%
rebellious	2.09	2.03	3.93	2.07	2.06	0.02%	0.06%	0.13%	0.12%
reckless	1.00	0.79	4.49	1.88	1.22	0.10%	0.07%	*0.21%*	0.11%
refined	4.37	4.42	2.15	3.39	4.06	0.08%	*0.29%*	0.04%	*0.24%*
religious	2.90	2.61	2.64	3.46	2.99	*1.02%*	*0.70%*	0.12%	*0.17%*
respectable	4.67	5.09	3.18	4.46	4.74	*1.01%*	*1.95%*	*0.15%*	*0.69%*
responsible	5.18	5.46	2.94	4.09	4.91	*0.49%*	0.15%	0.10%	0.06%
restless	2.29	1.46	3.46	2.51	2.09	0.16%	0.15%	*0.17%*	0.11%

righteous	3.52	3.61	3.36	3.69	3.61	*0.41%*	0.08%	0.01%	0.01%
romantic	3.27	4.51	1.46	5.04	4.27	0.09%	0.15%	*0.28%*	*0.53%*
rude	1.56	0.42	3.84	0.71	0.90	0.16%	0.08%	*0.42%*	0.12%
sad	1.99	0.84	2.60	1.43	1.42	*0.41%*	*0.49%*	*0.45%*	*0.37%*
selfish	1.78	0.79	3.39	0.97	1.18	*0.37%*	*0.62%*	*0.24%*	*0.27%*
sensible	4.82	5.04	2.50	3.94	4.60	*1.43%*	*2.21%*	*0.45%*	*2.20%*
sensitive	3.41	4.05	1.15	4.33	3.93	*0.66%*	*0.57%*	*1.19%*	*0.68%*
serious	4.27	3.89	3.67	2.40	3.52	*0.68%*	*0.30%*	*0.57%*	*0.46%*
shrewd	2.94	2.62	3.15	1.81	2.46	*0.66%*	*0.55%*	*0.15%*	*0.14%*
shy	2.46	2.23	1.73	2.80	2.50	*0.56%*	*0.37%*	*1.38%*	*1.48%*
silent	2.72	2.32	2.91	2.05	2.36	*1.04%*	*0.79%*	*0.68%*	*0.64%*
silly	2.47	3.11	2.77	4.09	3.22	*0.36%*	*1.72%*	*2.98%*	*3.12%*
smart	5.41	5.48	3.17	3.72	4.87	*0.78%*	*0.98%*	*2.04%*	*1.43%*
solemn	2.96	2.52	3.21	2.48	2.65	0.17%	0.04%	*0.29%*	*0.13%*
sophisticated	4.72	4.40	2.21	3.48	4.20	0.12%	*0.35%*	0.03%	0.11%
spirited	4.09	4.79	2.62	4.36	4.41	0.13%	*0.35%*	*0.39%*	*0.71%*
stern	3.43	2.37	4.26	1.59	2.46	*0.36%*	*0.21%*	0.02%	0.01%
stubborn	2.20	1.39	4.14	1.81	1.80	*0.32%*	*0.32%*	*0.23%*	*0.21%*
studious	4.94	4.54	2.83	3.37	4.28	0.09%	0.01%	*0.23%*	0.08%
talented	5.10	5.25	3.10	4.00	4.78	0.19%	*0.21%*	*0.18%*	*0.19%*
thoughtful	4.52	5.42	2.08	4.79	4.91	*0.36%*	0.19%	*0.30%*	*0.21%*
thoughtless	1.22	0.57	3.79	1.53	1.11	0.04%	0.03%	*0.20%*	*0.19%*
timid	2.10	1.64	1.74	2.49	2.08	*0.33%*	*0.41%*	*0.52%*	*0.61%*
tough	3.91	4.10	4.98	2.28	3.43	*0.31%*	*0.22%*	*0.43%*	*0.19%*
troubled	1.47	0.79	3.27	1.78	1.35	0.12%	0.12%	*0.15%*	0.11%
troublesome	1.37	0.61	3.84	1.51	1.16	0.02%	0.07%	*0.17%*	0.05%
ungrateful	1.27	0.57	3.38	1.05	0.96	0.09%	0.12%	*0.23%*	*0.32%*
unhappy	1.94	0.55	2.87	1.37	1.29	*1.30%*	*1.93%*	*1.01%*	*1.52%*
unlucky	1.96	0.75	3.21	2.02	1.58	0.14%	0.07%	*0.18%*	*0.13%*
unsophisticated	1.53	1.26	3.68	2.21	1.67	0.03%	0.03%	0.10%	*0.21%*
upright	4.08	4.49	3.28	3.61	4.06	*0.44%*	0.11%	0.06%	0.03%
vain	1.68	0.82	2.38	1.37	1.29	*0.38%*	*0.44%*	0.12%	*0.21%*
vigorous	4.01	4.06	3.97	3.56	3.88	*0.31%*	*0.29%*	0.10%	0.10%
vivacious	3.52	4.07	2.17	4.07	3.89	0.02%	*0.22%*	0.02%	*0.29%*
vulgar	1.51	0.89	4.73	1.48	1.29	0.14%	*0.45%*	*0.13%*	*0.14%*
warm	4.09	5.24	1.73	5.51	4.95	0.16%	*0.43%*	0.06%	*0.20%*
weak	1.24	0.84	1.61	2.17	1.42	*1.01%*	*1.32%*	*0.38%*	*0.49%*
wholesome	3.90	4.58	2.29	4.93	4.47	0.02%	0.11%	0.08%	*0.16%*
wise	5.34	5.42	3.54	4.09	4.95	*4.49%*	*3.22%*	*0.38%*	*0.34%*

Table 2. Ratings of adjectives and proportions of bigrams containing the adjectives. The table shows the average adjective ratings on four dimensions obtained in the present study, and the composite positivity score as well as proportions of bigrams containing the adjectives for each noun. The proportions of adjectives which are among the 100 most frequent adjectives for each noun are shown in italics.