



Parents mention sons more often than daughters on social media

Elizaveta Sivak^{a,1} and Ivan Smirnov^{a,1,2}

^aInstitute of Education, National Research University Higher School of Economics, Moscow 101000, Russia

Edited by Laura K. Nelson, Northeastern University, Boston, MA, and accepted by Editorial Board Member Mary C. Waters December 21, 2018 (received for review March 29, 2018)

Gender inequality starts early in life. Parents tend to prefer boys over girls, which is manifested in reproductive behavior, marital life, and parents' pastimes and investments in their children. While social media and sharing information about children (so-called "sharenting") have become an integral part of parenthood, whether and how gender preference shapes the online behavior of users are not well known. In this paper we use public posts made by 635,665 users from Saint Petersburg on a popular Russian social networking site, to investigate public mentions of daughters and sons on social media. We find that both men and women mention sons more often than daughters in their posts. We also find that posts featuring sons receive more "likes" on average. Our results indicate that girls are underrepresented in parents' digital narratives about their children, in a country with an above-average ranking on gender parity. This gender imbalance may send a message that girls are less important than boys or that they deserve less attention, thus reinforcing gender inequality from an early age.

gender inequality | son preference | parenthood | sharenting | social media

Gender inequality starts even before birth. Across the world, would-be parents tend to prefer their first (or their only) child to be a boy rather than a girl or to have more sons than daughters (1–8). This results in millions of “missing girls” at birth due to sex-selective abortions (9–11). Gender preference continues to manifest throughout childhood. In some countries, couples pursue sons by having additional children at the cost of having a larger family size and underinvesting in daughters (12, 13). Sons have advantages in nutrition (14), vaccination rates (15), and spending on healthcare (16, 17). Fathers (18–20) [and, in some cases, both parents (21)] spend more time with sons than with daughters. Fathers more often marry and stay married in families with sons (22, 23), although evidence for this is mixed (24, 25). Parents also report more happiness in families with sons (26).

Despite the extensive literature on gender preference, a study to examine whether the use of social media by parents is gender biased is needed. As social media become an integral part of parents' lives, understanding whether and how gender preference manifests in this environment is important. One common practice that has recently become a widespread trend is “sharenting” (27, 28), or parents' habitual use of social media to communicate detailed information about their children (sharenting, *Collins English Dictionary*, <https://www.collinsdictionary.com/dictionary/english/sharenting>). In this paper, we investigate gender preference in sharenting, drawing on data from 62 million public posts on a popular social networking site.

We obtained data from VK (vk.com), a Russian analogue of Facebook and the largest social networking site in Europe. VK provides an application programming interface (API) that allows for the systematic downloading of publicly available information. We used the VK API to collect public posts made in 2016 by 635,665 users from Saint Petersburg (the fourth largest European city), aged 18–50 y (see *SI Appendix* for details

on our sample and data collection). We then identified posts with mentions of children by examining posts that contained the words “daughter” and “son,” along with their different forms, e.g., “dochenka” (daughterling) or “sooooooon” (see *Materials and Methods* for details). Common topics for such posts included celebrations of different achievements and important events (e.g., births and birthdays or starting and finishing school); expression of love, affection, and pride; and reports on spending time with the children (see Fig. 1 for illustrative examples and *SI Appendix* for more information about common topics).

We computed the proportion of female and male users from each cohort who mentioned sons or daughters in their posts at least once, along with the average number of mentions of children for these users. In our analysis, we used various definitions for “mentions of children” to ensure that the results were not influenced by a specific choice of words (*Materials and Methods*). We also collected information about the number of “likes” that posts featuring children obtained on average. We used these data to investigate whether the social network environment might reinforce gender bias by rewarding posts featuring children of one gender more than those of another gender.

Results

Fig. 2 shows the proportion of users who mentioned children in their public posts at least once in 2016. The proportion of women

Significance

Parents' preference for sons is a well-known phenomenon. This study examines whether the use of social media by parents is gender biased. Due to the large-scale use of social media, even a moderate bias might significantly contribute to gender inequality. We use data from a Russian social networking site on posts made by 635,665 users and find that parents mention sons more often than daughters and that posts featuring sons get more "likes." This gender imbalance may send a message that girls are less important than boys or that they deserve less attention. Particularly in a country with an above-average ranking on gender parity, this invisible bias might present an intractable obstacle to gender equality.

Author contributions: E.S. and I.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. L.K.N. is a guest editor invited by the Editorial Board.

Published under the [PNAS license](https://www.pnas.org/licenses).

Data deposition: The data reported in this paper have been deposited at <https://osf.io/4ncbu/>.

¹E.S. and I.S. contributed equally to this work.

²To whom correspondence should be addressed. Email: ibsmirnov@hse.ru.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1804996116/-DCSupplemental.

Published online January 22, 2019.

- Michail Ivanov 29 Apr at 08.05 ...
 Daughter! 3500 g, 51 cm!!!
- Maria Ermolaeva 10 Apr at 11.56 ...
 10th of April... And only my cheerful little son was delighted with the snow
- Konstantin Leonidov 4 Feb at 18.14 ...
 Made a snowman with my daughter!
- Olga Stepanova 13 July at 20.06 ...
 The first medal! You are the best, son

Fig. 1. Selected examples of posts with mentions of children. All names and dates have been changed.

increases sharply until 31–32 y old and then gradually falls. The peak matches the average age of women at first childbirth, which is 30 y in Saint Petersburg (29). The proportion of men who mention children is significantly lower and steadily increases with age. In almost all cohorts of users, sons are mentioned by a larger proportion of both men and women. This difference cannot be explained by the sex ratio at birth alone (106 boys to 100 girls in Russia), thus indicating gender preference in sharing information about children. The exact estimate of the observed bias depends on the chosen measure and the set of words that are considered synonyms for the words son and daughter (see *SI Appendix* for detailed analysis).

The method that we use is not without limitations as it allows for false positive and false negative classification of the posts. For instance, “Gazprom’s daughter company” would be counted as a mention of a daughter and thus represents a false positive. Alternatively, people might refer to their daughters and sons using the generic term “child,” and such mentions would not be counted (false negative). If our method produces more false positives related to sons or more false negatives related to daughters, this might explain the observed bias.

To rule out these explanations, we randomly selected 10,000 posts that mentioned daughters and sons and manually checked whether they were about children. We found that 85% of them

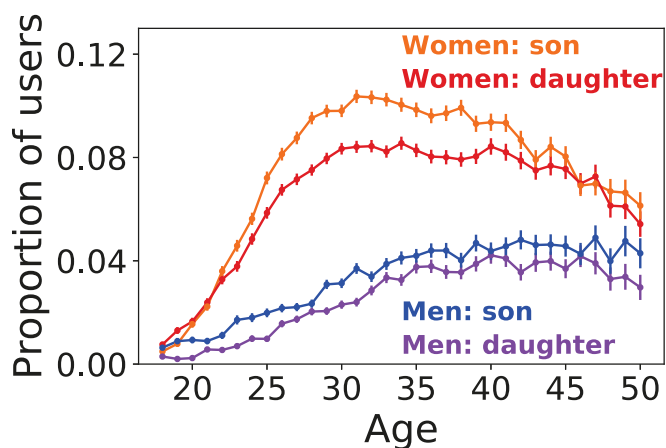


Fig. 2. The proportion of users who mentioned children in their public posts at least once in 2016. Sons are mentioned by a larger proportion of both men and women. Vertical bars are standard errors.

were about parents’ own children (detailed information is provided in *SI Appendix*, Tables S1 and S2). We used these data to compute the statistical significance of the difference between the number of posts about parents’ own daughters and sons, using the bootstrap test. We found that more posts about sons existed with $P < 10^{-4}$. We also performed an additional analysis and found that the bias persisted after the inclusion of generic references to children (*Materials and Methods*).

We also found that posts featuring sons were more rewarded; that is, they got more likes than those featuring daughters. Average numbers of likes are presented in Table 1. Here, three patterns can be distinguished. First, women “liked” posts more often than men. Second, a gender homophily in likes existed; i.e., women preferred posts written by women, and men preferred those written by men. Third, both women and men more often liked posts that mentioned sons.

Discussion

Studies of gender preference in parental practices usually have to rely on self-reports, e.g., reports about time spent with children (18–21). Self-report studies have some benefits, but their results are affected by various biases, including social desirability bias or recall bias. Mentions in posts are directly observable and present a clear and simple metric that can be used on easily accessible data to measure parents’ gender bias. We used this metric on a large dataset of public posts of more than 600,000 users and found that both men and women exhibited son preference on the social networking site: Sons were mentioned significantly more often than daughters. This result was remarkably stable and held true across age cohorts, different measures, and sets of words. We also found that posts in which sons were mentioned were more rewarded: These posts got around 1.5 times more likes than stories featuring daughters.

Son preference in traditional societies and developing countries is a well-known phenomenon. Our results confirm that son preference is also prevalent in countries not immediately associated with gender disparity. [Russia is above average in the ranking of countries by gender parity (30).]

Gender preference in sharenting may seem quite harmless compared with such layers of gender disparity as sex-selective abortions or underinvestment in girls. However, son bias online may contribute to daughters feeling underappreciated and less visible. It may also have broader effects as gender inequalities in everyday social interactions could translate to larger structures of inequality, leading to gender inequality even in advanced societies (31). Given the widespread popularity of social media, even moderate bias might accumulate. Son preference in likes can additionally amplify this bias, acting as social media’s built-in positive feedback loop. Millions of users are exposed to a gender-biased newsfeed on a daily basis and, without even noticing, receive the reaffirmation that paying more attention to sons is normal.

Previous studies have shown that children’s books are dominated by male central characters (32, 33). In textbooks, females

Table 1. The average number of likes per post

Posts	Average no. of likes	
	By women	By men
Written by women		
Featuring daughters	6.7	1.1
Featuring sons	10.7	1.8
Written by men		
Featuring daughters	5.3	2.6
Featuring sons	6.7	3.7

All daughters/sons differences are statistically significant with $P < 10^{-4}$.

are given fewer lines of text, have fewer named characters, and have fewer mentions than men (34). Additionally, in movies, on average, twice as many male characters as female ones are in front of the camera (35). While female coverage on Wikipedia compares favorably with that on some other lists of notable people (36), still, four times more articles about men than women exist (37). Gender imbalance in public posts may send yet another message that girls are less important and interesting than boys and deserve less attention, thus presenting an invisible obstacle to gender equality.

Materials and Methods

We used the API of VK to download all public posts of users from Saint Petersburg that were made in 2016 (available at ref. 38). We then computed vector representations of Russian words by training a fastText (39) model on the collected corpus (our model is available at ref. 38). We used this model to identify words similar to son and daughter, namely the closest words in the vector space measured by cosine distance. We manually excluded unrelated words. For instance, both the words son and “granddaughter” are unsurprisingly semantically close to the word daughter, according to the model. However, these are not synonyms for the word daughter, and we do not treat them as mentions of daughters.

After the exclusion of unrelated words, we obtained a list of the 30 closest synonyms to the word daughter and the 30 closest synonyms to the word son. Posts that included at least one of these words were considered to be posts mentioning children. The use of word embeddings trained on the VK corpus allowed us to consider words or their forms that cannot be found in dictionaries but that are used by the users of the social network, e.g., soon instead of son. We performed an additional analysis to ensure that our results were not driven by a particular choice of words (*SI Appendix*).

- Tian X, Yu X, Klasen S (2018) Gender discrimination in China revisited: A perspective from family welfare. *J Chin Econ Bus Stud* 16:95–115.
- Hank K, Kohler HP (2000) Gender preferences for children in Europe: Empirical results from 17 FFS countries. *Demographic Res* 2:1.
- Rossi P, Rouanet L (2015) Gender preferences in Africa: A comparative analysis of fertility choices. *World Dev* 72:326–345.
- Abeykoon A (1995) Sex preference in South Asia: Sri Lanka an outlier. *Asia-Pacific Popul J* 10:5–16.
- Bongaarts J (2013) The implementation of preferences for male offspring. *Popul Dev Rev* 39:185–208.
- Dahl E, Beutel M, Brosig B, Hinsch KD (2003) Preconception sex selection for non-medical reasons: A representative survey from Germany. *Hum Reprod* 18:2231–2234.
- Dahl E, et al. (2006) Preconception sex selection demand and preferences in the United States. *Fertil Steril* 85:468–473.
- Raley S, Bianchi S (2006) Sons, daughters, and family processes: Does gender of children matter? *Annu Rev Sociol* 32:401–421.
- Duthé G, Meslé F, Vallin J, Badurashvili I, Kuyumjian K (2012) High sex ratios at birth in the Caucasus: Modern technology to satisfy old desires. *Popul Dev Rev* 38:487–501.
- Miller BD (2001) Female-selective abortion in Asia: Patterns, policies, and debates. *Am Anthropologist* 103:1083–1095.
- The World Bank (2012) *World Development Report 2012: Gender Equality and Development* (The World Bank, Washington, DC).
- Zaidi B, Morgan SP (2016) In the pursuit of sons: Additional births or sex-selective abortion in Pakistan? *Popul Dev Rev* 42:693–710.
- Altindag O (2016) Son preference, fertility decline, and the nonmissing girls of Turkey. *Demography* 53:541–566.
- Barcellos SH, Carvalho LS, Lleras-Muney A (2014) Child gender and parental investments in India: Are boys and girls treated differently? *Am Econ J Appl Econ* 6:157–189.
- Borooh VK (2004) Gender bias among children in India in their diet and immunisation against disease. *Soc Sci Med* 58:1719–1731.
- Song L (2008) In search of gender bias in household resource allocation in rural China (IZA, Bonn), Discussion Paper 3464.
- Ganatra B, Hirve S (1994) Male bias in health care utilization for under-fives in a rural community in western India. *Bull World Health Organ* 72:101–104.
- Baker M, Milligan K (2016) Boy-girl differences in parental time investments: Evidence from three countries. *J Hum Cap* 10:399–441.
- Harris KM, Morgan SP (1991) Fathers, sons, and daughters: Differential paternal involvement in parenting. *J Marriage Fam* 53:531–544.
- Aldous J, Mulligan GM, Bjarnason T (1998) Fathering over time: What makes the difference? *J Marriage Fam* 60:809–820.
- Lundberg SJ (2005) The division of labor by new parents: Does child gender matter? (IZA, Bonn), Discussion Paper 1787.

We also removed potentially fake accounts and filtered posts that were not made by users themselves (see *SI Appendix* for details on data preprocessing) and then computed the proportion of users who mentioned children at least once in their posts as well as the average number of such mentions per user.

We used the same model to identify words other than daughter or son that might be used as a reference to children, e.g., “my little one.” Note that unlike in English, in Russian, most of these words have distinct masculine and feminine forms. This allows one to definitively identify whether they are mentions of daughters or sons. The inclusion of these additional words in the lists of synonyms did not affect the results. We found that among the words where gender could not be inferred from the word form, only one instance was frequent enough to have any effect on our results: the word “child,” along with its different forms. One way to check whether this was disproportionately used to refer to a daughter is to infer the gender of a child by looking at other posts made by the same user. We found that the reference to a child was accompanied by the reference to a daughter for 25.9% of users. This means that among those who used the word child at least once in their 2016 posts, 25.9% also used the word daughter or one of its closest forms at least once. The reference to a child was accompanied by the reference to a son for 28.9% of users and was not accompanied by daughter and son references for 54.3% of users. This result does not support the assumption that people may more often refer to the daughters as to the generic child compared with sons. Note that the sum is larger than 100% as some users mention both sons and daughters at some point in time. The extended list of words that might be used as a reference to children is available at ref. 38.

ACKNOWLEDGMENTS. Support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged.

- Dahl GB, Moretti E (2008) The demand for sons. *Rev Econ Stud* 75:1085–1120.
- Blau FD, Kahn LM, Brummund P, Cook J, Larson-Koester M (2017) Is there still son preference in the United States? (National Bureau of Economic Research, Cambridge, MA), Technical Report 23816.
- Lundberg S, McLanahan S, Rose E (2007) Child gender and father involvement in fragile families. *Demography* 44:79–92.
- Diekmann A, Schmidheiny K (2004) Do parents of girls have a higher risk of divorce? An eighteen-country study. *J Marriage Fam* 66:651–660.
- Kohler HP, Behrman JR, Skytthe A (2005) Partner + children = happiness? The effects of partnerships and fertility on well-being. *Popul Dev Rev* 31:407–445.
- Blum-Ross A, Livingstone S (2017) “Sharenting,” parent blogging, and the boundaries of the digital self. *Popular Commun* 15:110–125.
- Brosch A (2016) When the child is born into the Internet: Sharenting as a growing trend among parents on Facebook. *New Educ Rev* 43:225–235.
- Interfax (2017) The highest average age of women at first childbirth in Russia was recorded in Saint Petersburg. Available at www.interfax.ru/russia/570614. Accessed March 21, 2018.
- The World Economic Forum (2016) *Global Gender Gap Report 2016* (The World Economic Forum, Geneva).
- Ridgeway CL (2011) *Framed by Gender: How Gender Inequality Persists in the Modern World* (Oxford Univ Press, Oxford).
- McCabe J, Fairchild E, Grauerholz L, Pescosolido BA, Tope D (2011) Gender in twentieth-century children’s books: Patterns of disparity in titles and central characters. *GenD Soc* 25:197–226.
- Hamilton MC, Anderson D, Broadus M, Young K (2006) Gender stereotyping and under-representation of female characters in 200 popular children’s picture books: A twenty-first century update. *Sex Roles* 55:757–765.
- Blumberg RL (2008) The invisible obstacle to educational equality: Gender bias in textbooks. *Prospects* 38:345–361.
- Smith SL, et al. (2015) *Inequality in 700 Popular Films: Examining Portrayals of Gender, Race & LGBT Status from 2007 to 2014* (University of Southern California Annenberg School for Communication and Journalism, Los Angeles).
- Wagner C, Garcia D, Jadidi M, Strohmaier M (2015) It’s a man’s Wikipedia? Assessing gender inequality in an online encyclopedia. *Proceedings of the 9th International Conference on Web and Social Media*, eds Quercia D, Hogan B (AAAI Press, Palo Alto, CA), pp 454–463.
- Swedish Ministry for Foreign Affairs (2017) WikiGap. Available at <https://meta.wikimedia.org/wiki/WikiGap>. Accessed March 21, 2018.
- Sivak E, Smirnov I (2018) VK text. Available at <https://osf.io/4ncbu/>. Deposited August 2, 2018.
- Bojanowski P, Grave E, Joulin A, Mikolov T (2016) Enriching word vectors with subword information. arXiv:1607.04660 Preprint, posted June 19, 2017.