

Sexist Textbooks

 Lee Crawford, Christelle Saintis-Miller, Rory Todd

ABSTRACT

Textbooks play a critical role in schooling around the world. Small sample studies show that many books continue to under-represent women and girls, and to portray men and women in stereotypical gendered roles. In this paper, we use quantitative text analysis to assess the degree of gender bias in a newly assembled corpus of 1255 English language school textbooks from 34 countries that are publicly available online. We find consistent patterns of under-representation of female characters and portrayal of stereotypical gendered roles. Women and girls occur less frequently, are portrayed as more passive, are less likely to be associated with work or achievement, and are more likely to be associated with the home and traditionally female occupations. Comparing across countries, female representation in books is correlated with higher GDP and more legal rights for women.

KEYWORDS

Gender, Education,
Textbooks

JEL CODE

J16

Sexist Textbooks

Lee Crawford, Christelle Saintis-Miller, and Rory Todd

Center for Global Development

We're grateful to Laura Moscoviz and Emma Cameron for research assistance, to Anjali Adukia, Susannah Hares, Alice Evans, Margaret Leighton, David Evans, Asyia Kazmi, Izzy Boggild-Jones, and Shubha Jayaram for helpful comments and advice, and to Ezioma Akabike and Mo Olateju for assistance in sourcing textbooks. This research was supported by funding from Echidna Giving and the Bill & Melinda Gates Foundation.

If you would like to access the original educational materials or replicate the process of converting the original PDF files to TXT files, download the folder available [here](#). Please note that the folder is very large (27GB). The data and code are available [here](#).

Lee Crawford, Christelle Saintis-Miller, Rory Todd. 2024. "Sexist Textbooks." CGD Working Paper 678. Washington, DC: Center for Global Development. <https://www.cgdev.org/publication/sexist-textbooks>

CENTER FOR GLOBAL DEVELOPMENT

2055 L Street, NW Fifth Floor

Washington, DC 20036

202.416.4000

1 Abbey Gardens

Great College Street

London

SW1P 3SE

www.cgdev.org

The Center for Global Development works to reduce global poverty and improve lives through innovative economic research that drives better policy and practice by the world's top decision makers. Use and dissemination of this Working Paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The views expressed in CGD Working Papers are those of the authors and should not be attributed to the board of directors, funders of the Center for Global Development, or the authors' respective organizations.

1. Introduction

School textbooks play an important role in shaping identity, attitudes, (Cantoni et al., 2017; Cantoni and Yuchtman, 2013; Chen et al., 2023; Lee, 2023; Miguel, 2004), and norms (Dhar et al., 2022; Mehmood et al., 2022). Controversy over school book content is high in many countries including the US, where book bans continue today. The introduction of secular textbooks was critical in the political development of the religious right in America, and led to school boycotts, a shooting, and a bombing (Kincheloe, 1983). Despite the potential importance of books for norms, still little is known quantitatively and systematically about the extent of biases in school books around the world.

In this paper we open the black box of global schooling, and provide quantitative insight into the content of school curricula and learning materials. Our particular focus is on girls education and gender bias. Girls education is a major global policy priority, but attention has focused primarily on the amount of schooling girls receive and the cognitive skills they obtain. If we care about equality of opportunity for boys and girls, we need to look beyond the building of human capital, to what children are exposed to in schools. Though there are many influences on children’s gender norms, school books make up a large share of children’s time spent at school, and may pass on implicit biases (Witt, 2001). Many teachers make a majority of their instructional decisions based on textbooks (Sadker and Zittleman, 2007).

Concretely, in this paper we ask how representation of males and females quantitatively differs across English-language textbooks globally, and to what extent textbooks portray stereotypes concerning the occupations, characteristics, and actions of male and female characters.¹

To do this we use a newly collected corpus of text documents, covering 1255 school textbooks for 34 countries. We process these documents using modern natural language processing and quantitative text analysis tools. We identify differences in the quantitative representation of male and female coded words, as well as in the qualitative descriptions of male and female words, using word co-occurrence, word embeddings, and part-of-speech tagging.

Overall we find substantial gender bias in school textbooks across the 33 countries, with women and girls under-represented compared to men and boys, and portrayed in stereotypically feminine roles. In 28 of the 34 countries, women and girls occur less frequently than men and boys, and the gap for some countries is stark. Women tend to be portrayed in traditionally feminine occupations, such as nurse, teacher, and housekeeper. They are more likely than men to be associated with words relating to appearances

¹We use ‘textbooks’ throughout as shorthand, though a small number of materials in our corpus are in fact student worksheets or slideshows.

and the home, compared to achievement and work. Women and girls are described as being more passive and less dominant than men, although they are also portrayed using more positive terms.

Comparisons between countries are limited by small samples, but higher income countries tend to have somewhat less unequal representation, although stereotypes still seem to be common. Countries with better representation also have better legal rights for women, even after adjusting for income levels.

Our analysis complements existing studies that have documented consistent gender bias in textbooks around the world, primarily through manual analysis and coding of texts (Blumberg, 2008; Mustapha and Mills, 2015). Studies on gender representation in textbooks from low and middle-income countries come primarily from small-scale qualitative manual coding of books. Such studies typically show an under-representation of female characters, and associate women with lower-status occupations than men. This includes studies from Bangladesh (Asadullah et al., 2018), Cameroon (Brugeilles and Cromer, 2009), China (Ye, 2022), Ethiopia (Bachore and Semela, 2022), Hong Kong (Lee and Collins, 2010), Iran (Bahman, 2020), Nepal (Bhattarai, 2020), Nigeria (Mustapha, 2014), Pakistan (Mirza, 2004), and Palestine (Karama, 2020). Some studies also present comparisons across a small number of countries, such as Bangladesh, Indonesia, Malaysia, and Pakistan (Islam and Asadullah, 2018), Australia, Hong Kong, Singapore, and Turkey (Lee and Collins, 2010; İncikabi and Ulusoy, 2019), and Rwanda, Kenya, and Uganda (Barton and Sakwa, 2012; Russell et al., 2021).

A key advantage of our study is the ability to compare across a wider range of countries. Such larger comparative cross-country research is rare. One exception is a project at Stanford University looking at changes over decades, using books from the Georg Eckert Institute for International Textbook Research. Jimenez et al. (2017) review over 1,000 secondary school textbooks from 88 countries from between 1950 and 2011. They show an increase in mentions of women’s rights over time. Buckner and Russell (2013) code 559 secondary school textbooks from 76 countries. Their focus is on representations of globalisation rather than gender, and rely on manual coding of books. Skinner and Bromley (2019) use a similar approach to look at long-run trends in discussion of human rights in textbooks.

Methodologically our study is similar to some others applying natural language processing to school textbooks from high-income countries. For example Adukia et al. (2023) analyse the representation of race and gender in 1,000 award-winning American children’s books, using both text and image analysis techniques. Adukia et al. (2022) study word embeddings in the same corpus, finding that females are more likely to be represented in relation to their appearance than in relation to their competence, and more likely to be represented in relation to their role in the family than their role in business. Lewis et al. (2022) study the best-selling 247 books aimed at under 5 year olds in the US and Canada, finding that many books include gender stereotypes such as that girls are better at reading and boys at math. Lucy et al. (2020) use word embeddings to show that in American history textbooks, women are discussed more

in the context of the home and less in the context of achievement. More broadly, the use of quantitative text analysis to document gender bias is growing across a range of topics in economics and political science, including in judge’s opinions (Ash et al., 2024), Google News and Google Books (Bolukbasi et al., 2016; Garg et al., 2018), local newspapers (Singla and Mukhopadhyay, 2022), and the Corpus of Historical American English (Boutyline et al., 2023).

Other studies on the impact of textbooks in developing countries have focused on the provision of textbooks rather on their specific content. For instance Glewwe et al. (2009) randomise access to new books in Kenya, finding no impact on learning on average, and positive impacts only for the most able children. Similarly Kuecken and Valfort (2013) find that providing textbooks has zero effect on learning on average across 11 anglophone African, but positive effects for children from the wealthiest families, in this case using student fixed effects models. By contrast, packaged interventions including student books and teacher guides and training have been more successful, demonstrated experimentally in Kenya (Piper et al., 2018) and El Salvador (Maruyama and Kurosaki, 2022).

In the remainder of this paper we discuss the data we use in Section 2, the methods in Section 3, and our results in Section 4.

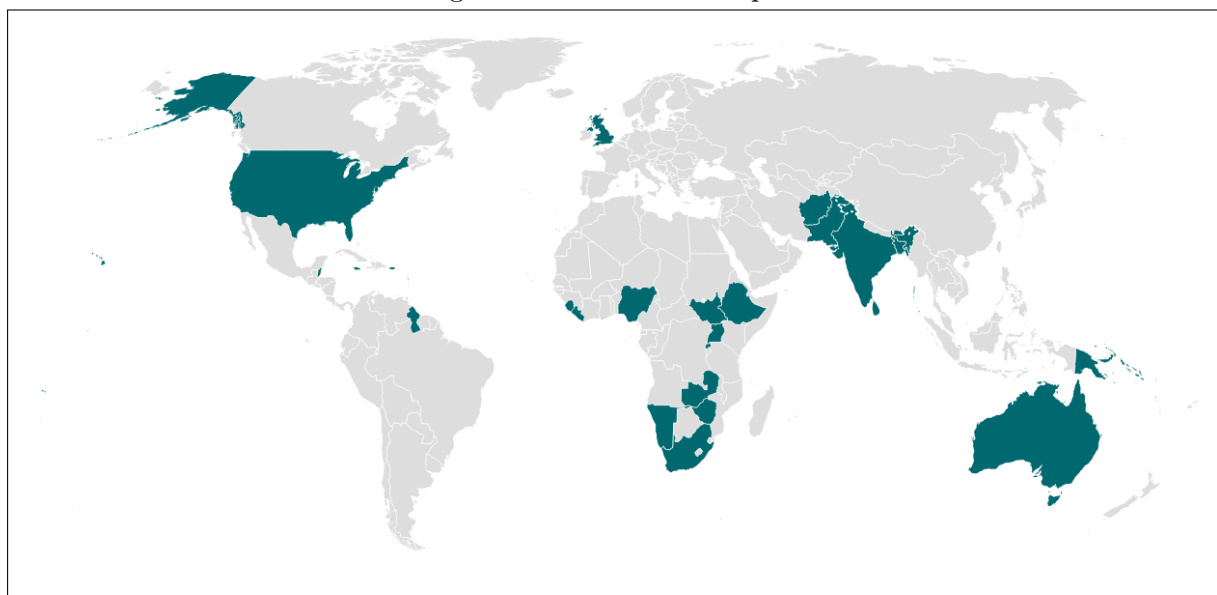
2. Data

2.1. Textbook Corpus

Our data comes from a systematic review of online materials. From September 2020 to February 2022, we visited every ministry of education website in the world and downloaded all available educational materials (many countries made learning materials publicly available online to facilitate online learning during the Covid-19 pandemic). After selecting only files that were in English and which could be classified as learning materials - rather than teacher guides or subject curricula - we found that our sample had a small range of countries, with a disproportionate contribution by small island countries. To ensure a broad range of countries were represented, we searched for publicly available learning materials from websites other than ministries of education, focusing on countries with sizable populations and where English is an official language. Almost all of these books are published by government bodies, with the exception of the three high-income countries in our sample, where textbooks are published exclusively by the private sector (Australia, the United Kingdom, and the United States).

Our sample contains 1255 textbooks from 34 countries, which total over 37.2 million words. These books cover a wide range of school subjects from grades 4 to 13. The included countries are: Afghanistan, Australia, Bangladesh, Belize, Bhutan, Dominica, Ethiopia, Guyana, India, Jamaica, Kenya, Kiribati, Lesotho, Liberia, Malawi, Maldives, Namibia, Nigeria, Pakistan, Papua New Guinea, Rwanda, Samoa, Sierra Leone, Solomon Islands, South Africa, South Sudan, Sri Lanka, St Kitts and Nevis, Tonga, Uganda, United Kingdom, United States, Zambia, and Zimbabwe (Figure 1).

Figure 1: Countries in corpus



Note: Books from the 34 countries included in our corpus are coloured in teal. This includes Afghanistan, Australia, Bangladesh, Belize, Bhutan, Dominica, Ethiopia, Guyana, India, Jamaica, Kenya, Kiribati, Lesotho, Liberia, Malawi, Maldives, Namibia, Nigeria, Pakistan, Papua New Guinea, Rwanda, Samoa, Sierra Leone, Solomon Islands, South Africa, South Sudan, Sri Lanka, St Kitts and Nevis, Tonga, Uganda, United Kingdom, United States, Zambia, and Zimbabwe

While data on actual use of textbooks is scarce, the data that is available points to textbooks as being widely available and playing a dominant role in the curricular resources used in classroom instruction. Whilst there has long been concern about a scarcity of textbooks in poorer countries (Fredriksen and Brar, 2015), the majority of pupils in school surveys do report access to at least a shared textbook. The 2019 PASEC survey of 14 Francophone African countries² found that on average 60-75 percent of 13-year olds had their own textbooks (Bietenbeck et al., 2023). The 2018 PISA for Development survey of 15-year olds in seven countries³ found that overall around 45 percent of students had their own textbook, and another 45 percent sharing with another student or students, with only 10 percent reporting no access at all (Ward, 2018). The most recent 2012 SACMEQ survey of 14 Anglophone African countries, found that around 40 percent of grade six pupils owned reading and mathematics textbooks (Awich, 2021), with that number roughly doubled through sharing (Kuecken and Valfort, 2013). A 2015 survey in Rwanda found that though textbooks were widely available in all schools, they were only observed to be present in around three out of five classrooms (Milligan et al., 2017).

Textbooks are also used as a primary instructional tool in the classroom, influencing not only what teachers teach, but how they teach (Stará et al., 2017). The 2011 Trends in International

²Benin, Burkina Faso, Burundi, Cameroon, Chad, Republic of the Congo, Democratic Republic of the Congo, Gabon, Guinea, Ivory Coast, Madagascar, Niger, Senegal, and Togo.

³Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal, and Zambia.

Mathematics and Science Study (TIMSS) included survey questions on the types of materials used as the basis of instruction. 70% of fourth grade science teachers and 74% of eighth grade science teachers used textbooks most often as the basis for instruction (Martin et al., 2012). In mathematics instruction, 75% of fourth grade teachers and 77% of eighth grade teaches used textbooks most frequently. The World Bank Service Delivery Indicator surveys conducted between 2012 and 2016 in 8 African countries found that 37 percent of children in randomly selected grade 4 classrooms had the relevant textbook (World Bank, 2023)⁴. A study observing classrooms in successful reading programs found that they spent large shares of class time using textbooks - 50 percent in Kenya, 45 percent in India, 42 percent in Senegal, 28 percent in Nigeria, and 8 percent in Tanzania (Stern et al., 2023).

In the majority of countries in our sample, the government provides a set of specific approved textbooks directly to schools, rather than schools making their own choices. Therefore particularly in these countries with centralised supply we can expect large proportions of children in school to be actually exposed to the books in our corpus (Read, 2015). Typically, poorer countries have a single set of government-approved and provided books, whereas in richer countries schools are afforded more autonomy to select books from a range of providers.⁵

2.2. Gendered Word Lists

We compare the content of the textbooks in our corpus with pre-defined lists of gendered nouns, pronouns, and names. We adapt the gendered nouns and pronouns used in Adukia et al. (2023), and include words such as she/her/woman and he/his/man. In total there are 68 male words and 68 female words (listed in Table C1).

To identify names and their gender in our text, we use *damegender* (Menendez, 2022), a tool for gender detection generated using census data from 20 countries. Given the international source of the textbooks, we find this performs better than datasets using only names from one particular country like the United States. We only allow names above a frequency of 10,000 across this dataset, as names with a frequency below this threshold were significantly less likely to be names in context. We exclude other names unlikely to be names in context, including country names, city names, and months of the year. We exclude names where the gender of the individual may be ambiguous, defined as those where neither gender accounts for at least 95% of occurrences in the dataset.

⁴This ranged from 88% in Morocco, through 68% in Togo, 66% in Mozambique, %43 in Kenya, 40% in Nigeria, 19% in Tanzania, 14% in Niger, to 11% in Madagascar

⁵Government-approved books are distributed to schools in 23 countries in our sample; Afghanistan, Bangladesh, Bhutan, Ethiopia, Kenya, Kiribati, Lesotho, Liberia, Maldives, Nigeria, Papua New Guinea, Rwanda, Samoa, Seychelles, Sierra Leone, Solomon Islands, South Sudan, Sri Lanka, Tonga, Trinidad and Tobago, Uganda, Zambia, and Zimbabwe. Schools have autonomy over textbook choice in Australia, Belize, Dominica, Guyana, Jamaica, Namibia, South Africa, St Kitts and Nevis, United Kingdom, and United States. India and Pakistan have a mix of central government-provision and local school choice in different states and provinces.

Table 1: Number of Books, by Country and Grade

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Afghanistan				1	1	1	1	1	1	1	1	1		9
Australia	10	8	7	4	6	7	11	7	11	7				78
Bangladesh				7	6	16	16	15		26				86
Bhutan				8	2	1	1	1		1	1	1		16
Dominica					2	2	1							5
Ethiopia		1	1		1	2	4	1	7	6	1	9		33
Guyana	11	5	5	5	4	4								34
India	3	4	5	9	9	12	12	12	19	20	26	32		163
Kenya	9	2	1											12
Kiribati				3	7	7	6		1				1	25
Malawi	1	1	1	1		1								5
Maldives	10	8	7											25
Namibia										5		1		6
Nigeria	9	2												11
Other	2			5	1				1	4	2	2		17
Pakistan	3	3	3	7	7	9	8	10	16	10	10	8		94
Papua New Guinea			3	3	3	3								12
Rwanda	5	7	5	4	4	3	3	2	4	4	4	5		50
Samoa									17	18	13	15	7	70
Sierra Leone							6	6	5	6	6	10		39
Solomon Islands				8	8	8	12	7	8	8				59
South Africa				7	6	6	9	8	6	3	2	2		49
South Sudan	5	5	5	5	5	5	4	5	10	10	10	11		80
Sri Lanka				1	1	9	24	10	10	27	9	1		92
St Kitts and Nevis				4	4	4	7	6	1	13	9	6		54
Uganda								16	2	2	3	3	1	27
United Kingdom		3	1	4	1	2	3	2	3			1		20
United States	7	2	7	9	9	7	6	5	8	4	4	4		72
Zimbabwe		3		3			2		4					12
Total	75	54	51	98	87	109	136	114	134	175	101	112	9	1,255

Note: This table shows the number of textbooks in our corpus in each grade in each country. “Other” includes books from Belize, Jamaica, Lesotho, Liberia, Tonga, and Zambia - we treat these separately as there are fewer than 5 books from each individual country.

Table 2: Number of Books, by Country and Subject

	Ag	Ar	Ec	Ge	Ho	Hu	IT	La	Ma	PE	Re	Sc	So	Total
Afghanistan								9						9
Australia								46	28			4		78
Bangladesh	4	4	4	9		10	4	10	7	4	21	9		86
Bhutan								10	3	1		1	1	16
Dominica									1			4		5
Ethiopia				1	4	2	3	10	3			10		33
Guyana					6			6	15			7		34
India		3	8	3	2	21	2	32	34	1		32	25	163
Kenya								9	3					12
Kiribati			4		1	3	1	3	5			8		25
Malawi				1				4						5
Maldives		1					3	3	5	1	1	3	8	25
Namibia			1			2		1				2		6
Nigeria								11						11
Other			1	5			2	3	4			2		17
Pakistan				9	2	3	10	20	12	1	8	27	2	94
Papua New Guinea									8			4		12
Rwanda						5		8	9		5	23		50
Samoa	6	8	10		8	3		10	1	2		14	8	70
Sierra Leone								20	19					39
Solomon Islands								32	24			3		59
South Africa				3			2	6	22			16		49
South Sudan				4		8	4	12	14		10	20	8	80
Sri Lanka			10	4		15	9	23		15	6	9	1	92
St Kitts and Nevis	1			9		9	6	6	7			13	3	54
Uganda	1	3	1		1	7	1	1	1	1	2	8		27
United Kingdom			1	1		5		2	9			2		20
United States			2	2		5		47				16		72
Zimbabwe						2		8	2					12
Total	12	19	43	60	14	100	47	352	236	26	53	237	56	1,255

Note: This table shows the number of textbooks in our corpus in each subject in each country. Ag = Agriculture, Ar = Arts and Music, Ec = Economics and Business, Ge = General Studies, Ho = Home Economics, Hu = Humanities, IT = Information Technology, La = Languages, Ma = Maths, PE = Physical Education, Re = Religion, Sc = Sciences, So = Social Science. "Other" includes books from Belize, Jamaica, Lesotho, Liberia, Tonga, and Zambia - we treat these separately as there are fewer than 5 books from each individual country.

We validate this process manually with a sample of 5 books across a range of subjects (English, Mathematics, Social Studies, Political Science, and Biology) and countries (Afghanistan, Bangladesh, Guyana, India and Pakistan), to calculate both the rate of precision - the proportion of identified 'names' which are in fact names in context - and recall - the proportion of names in the text which were identified. This validation process showed an average precision rate of 70% - meaning from the names the tool identified, 70% were actually names - and an average recall rate of 75% - meaning the tool correctly identified 75% of the total names across the texts. We also examined the precision and recall rates across male and female names. The precision rate for female names was 65% compared to 78% for male names, meaning the tool was able to identify male names that were actually names in the context of the text at a higher rate than female names. The tool was able to identify the number of both female and male names at a similar rate, with the recall rate for female names 74% compared to 77% for male names. Because of this imprecision, we don't include names when measuring representation, instead only counting gendered nouns and pronouns. We do include names in the other parts of our analysis.

We have equal length lists of 68 male nouns/pronouns and 68 female nouns/pronouns, but more unique female names (2,218) than male names (1,753).

2.3. Family and Work Word Lists

To construct lists of words in which we might expect gender bias, we adapt lists used by Adukia et al. (2022), Kumar et al. (2021), and Lucy et al. (2020). These include 13 words related to achievement, 6 related to appearance, 11 related to family and the home, and 20 related to work (full word lists are shown in Table C3).

To classify occupations we start with a list of 1,156 occupation titles drawn from the Gazette (this is the official public record in the UK, and a commonly used source of text data, (Sheridan, 2015)). We discard occupations which occur fewer than ten times in our full textbook corpus or which are explicitly gendered such as *salesman* or *actress*. We then manually classify these occupations into three groups; professional, service, and manual occupations. We are left with 56 professional or managerial jobs, 19 service jobs, and 16 manual jobs (Table C2).

2.4. Country-related gender characteristics

We use several sources of data on gender norms and outcomes at the country level. First, the 2023 Social Institutions and Gender Index (SIGI) from the Organization for Economic Cooperation and Development (OECD). The SIGI index comprises 25 indicators covering discrimination, violence, bodily

autonomy, economic rights, and civil liberties. Second, the World Bank Women, Business, and the Law Index measures women's economic and legal rights, and comprises 35 indicators. Third the Center for Global Development (CGD) Girl's Education Policy Index (GEPI) measures policy effort related to girl's education across 32 indicators, including spending, sexual health, safety, employment, and role models (Crawford and Hares, 2020). Fourth, to measure equality in education *outcomes* we use the World Bank Gender Parity Index in Gross Secondary Enrolment. Fifth, we look at data on teenage marriage of girls, and on female MPs, both drawn from the 2023 OECD Gender, Institutions, and Development (GID) Database. We standardise all variables to a mean of zero and standard deviation of one.

3. Methods

3.1. Data processing

We begin by converting PDF books to text data, primarily using Optical Character Recognition (OCR), specifically in Amazon Textract. This process uses a pre-trained model to infer text in an image, and recent advances in machine learning have greatly improved its accuracy; a recent benchmarking experiment by Hegghammer (2022) found a word error rate of just 1.8% when Textract was used to convert a clear image. We validate the conversion process by manually comparing a sub-sample of text files to the original PDFs. For several countries where textbooks had machine-readable text, we used a function in R which utilises the poppler library of PDF tools. This tool is able to extract text directly, and has comparable accuracy.

We group books by subject into 14 groups; Agriculture, Arts, Business and Economics, Environmental Studies, French, General and Social Studies, Home Economics, Humanities, IT, Language, Maths, Physical Education, Science, and Social Science.

3.2. Measuring gender representation

We then measure the quantitative representation of male and female words in the corpus of textbooks. We do this by counting gendered nouns and pronouns. For each book we calculate the share of female words as a percentage of all gendered words. We then show how the share of female words varies by country, subject, and grade. We also count the number of gendered words with a clear age classification (e.g., girl vs. woman and boy vs. man), and whether gendered words are capitalised or not, as an indication of whether characters are the active subject or passive object of a sentence, both following Adukia et al. (2023).

3.3. Measuring gender stereotypes

3.3.1. Co-occurrence

The simplest way of measuring stereotypes is counting the number of co-occurrences within a single sentence of gendered words and specific target words. We then count all co-occurrences of these occupations with gendered terms. For this analysis we assume that the male or female term refers to the individual with the occupation. Whilst this analysis is conceptually simple, this assumption may introduce some noise and bias. For example the sentence “*he went to the doctor*” would count as a co-occurrence of a

male term with the occupation doctor, though in this case the sentence does not indicate the gender of the doctor. We omit cases where both male and female terms are mentioned within the same sentence.

3.3.2. Word embedding

A more sophisticated approach to measuring stereotypes uses word embeddings. Word embeddings are numerical representations of words that capture their semantic and syntactic properties. We use the Word2vec algorithm (Mikolov et al., 2013), which encodes words as high-dimensional vectors, or embeddings.⁶ Words which are mathematically closer in this vector space tend to be used in similar contexts in the training corpus, and can therefore be regarded as more closely associated. Embeddings can also be combined arithmetically. An oft-cited example is that the vector resulting from $king - man + woman$ will in a well-trained model be very close to the embedding for *queen*.

Word embeddings can be used for various natural language processing tasks, such as text classification, sentiment analysis, and machine translation. But they also reflect human-like biases that are present in the text data on which they are trained. For example, gender bias can be observed in word embeddings when certain words are more associated with male or female attributes or occupations. In our case, we are interested in the distance between male and female words on the one hand, and words relating to work, family, achievement, and appearance on the other.

Word embeddings require a relatively large sample of words and sentences, and so we train our model on our entire corpus rather than on individual countries or sub-samples of the data. We first estimate the embedding for each individual word. Second, we calculate the cosine similarity between each gendered word and each target theme word. We do this only for gendered nouns and pronouns, rather than names: many names will be very low frequency and embeddings will be correspondingly imprecisely measured. The similarity can take values between -1 and +1. Identical vectors have a cosine similarity of 1, perpendicular vectors have a 90-degree angle between them and a cosine similarity of 0, and opposite vectors have an angle of 180 degrees between them and a cosine similarity of -1. Third, we take the average similarity between all male words and each target word, and all female words and each target word. Fourth, we calculate the difference between the male similarity and the female similarity for each target word. This leaves us with a measure of male bias for each individual home, work, achievement, or appearance-themed word. Fifth, in order to assess the uncertainty in this male bias measure, we bootstrap standard errors, following Lucy et al. (2020). We do this by splitting the entire corpus into sentences, sampling sentences with replacement, and training a new model on this new sample. We

⁶Technically, word2vec can be used to learn associations between any kind of token: for example, two-word phrases, or sentences. Word2vec trains a neural network on a corpus to perform one of two word association tasks: if implemented with a 'continuous bag-of-words model', to predict a word based on the words surrounding it; and if implemented with a 'skip-gram' model, to predict the words surrounding a given word.

repeat this process 50 times, and take the standard deviation of the pro-male bias estimates across the 50 models as the estimate of the standard error.

For our primary set of results, we use a standard continuous bag-of-words model, and following the advice of Rodriguez and Spirling (2021), select a window size of 6, and generate embeddings with 300 dimensions. We use negative sampling which reduces the computational cost of predicting embeddings by sampling a few negative examples – words that are not likely to appear in the context of the input word, and minimizing their probability. We calculate the similarity of embeddings as their cosine product.

3.3.3. Part-of-speech tagging

An alternative approach to word embeddings is part-of-speech tagging, which infers the relationships between words in a sentence. Concretely, the ‘universal dependencies’ framework (de Marneffe et al., 2021) allocates all words in a language to a set of categories (adjective, noun, verb, etc), depending both on the word itself and the context in which it used (for instance in English ‘butter’ might be a noun or used as a verb as in ‘to butter the bread’). We make use of the UDPipe model trained on the ‘English Web Treebank’ (Bies, Ann et al., 2012; Silveira et al., 2023), a corpus of 16,621 sentences sourced from the internet that have been manually annotated. Applying this model to our corpus allows us to tag words, and identify the verbs and adjectives which are associated with male and female words. We can thus see how male and female characters are depicted, what actions they are depicted doing, and how active or passive they are (for instance depending on whether they are the *subject* or *object* of sentences).

Finally, in order to compare quantitatively the differences in how genders are portrayed across different country income levels, we conduct a sentiment analysis of the terms used with each gender. Specifically, we use the National Research Council Canada (NRC) Valence, Arousal, and Dominance (VAD) lexicon (Mohammad, 2018) which quantifies three aspects of words: valence - how pleasant or positive a word is - arousal - how active it is - and dominance - how dominant or submissive it is. This lexicon includes human ratings for more than 20,000 words. We calculate the average rating for each of the three dimensions for verbs and adjectives used with male and females terms, weighted by the frequency of the word.

3.4. Cross-country correlations

Does the existence of bias in textbooks correlate with other measures of gender equality at the country level? We look at the cross-country correlations between bias in textbooks and various measures of gender equality. There are several important limitations to this analysis. First, we aren’t able to make causal

claims here, and causality could clearly run in both directions; from biased books to unequal norms, and vice-versa from unequal norms to biased books. Nonetheless we feel this is a useful test of the validity of the text analysis - a complete lack of correlation could raise doubts. Second, our sample of books is essentially a convenience sample, and is not necessarily nationally representative of the books actually read by children in classrooms each country. Third, we have a relatively small sample of countries at just 34, and so this analysis has low statistical power.

4. Results

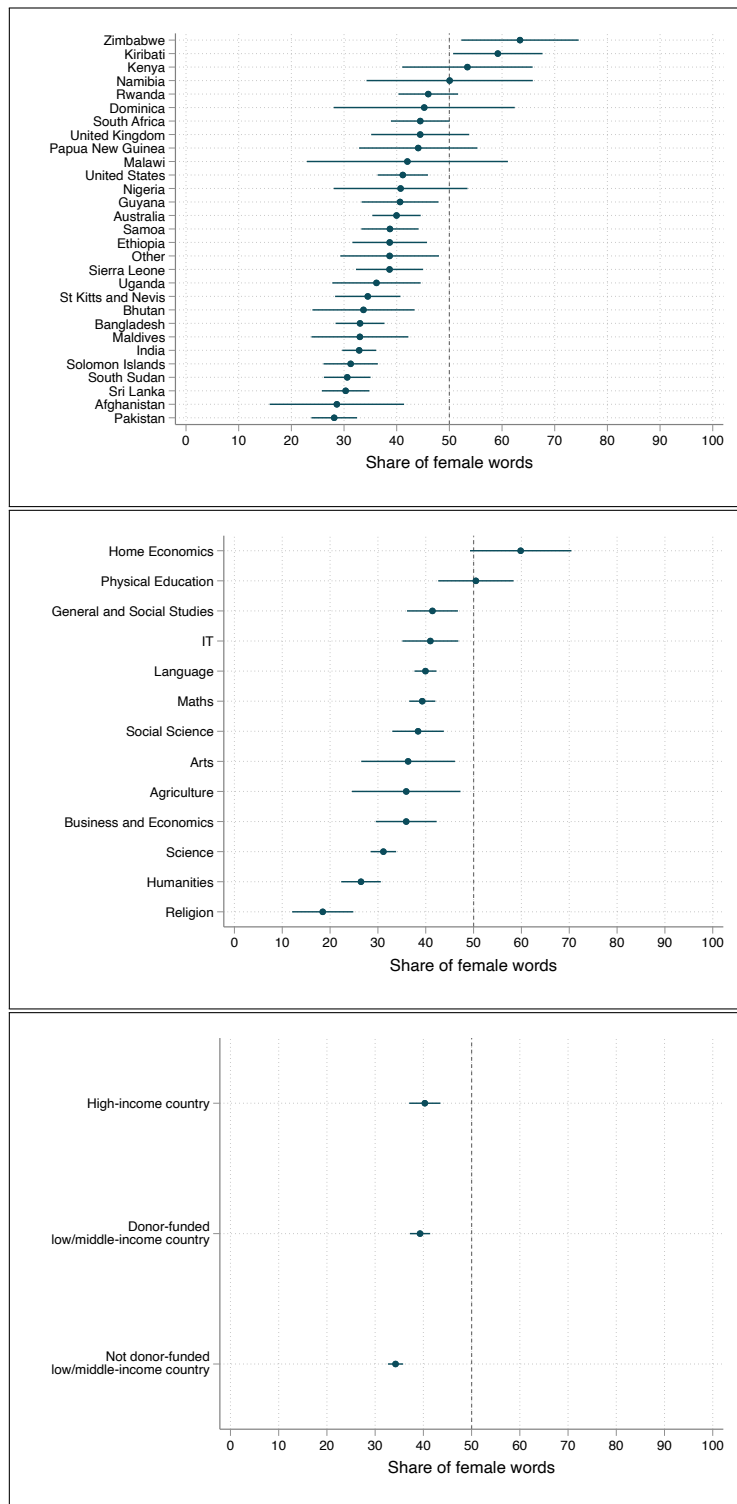
4.1. Gender representation

Overall across our full sample of books there are over twice as many occurrences of male words (178142) as there are female words (82113). There is also substantial variation across countries. After adjusting for book length, grade, and subject, the countries with the lowest representation of women and girls are Afghanistan, Pakistan, Sri Lanka, and South Sudan, where less than 1 in 3 gendered words are female (Figure 2). Inspecting difference across subjects (across all countries), the lowest representation of female words are in religious studies and the humanities. High-paying subjects such as science and maths also have less than equal representation.⁷ The subject with the highest female representation is home economics. Female representation decreases with book length, and there are no statistically significant differences in female representation across different grades. Amongst low- and middle-income countries, those with books funded by international aid projects have higher but still uneven female representation.

Overall we see more equal representation of young girls than adult women. Of gendered words with a clear age classification (e.g., girl vs. woman and boy vs. man), female words make up 47 percent of child words, but only 41 percent of adult words (this is similar to the pattern observed in the United States by Adukia et al. (2023)). Focusing on the share of gendered words that are capitalised, we see that 29.07 percent of male words are capitalised, compared to only 23.63 percent of female words. This indicates that male characters are more likely to be the subject of a sentence and therefore taking a position of active importance, rather than being in a more passive role (Figure 3).

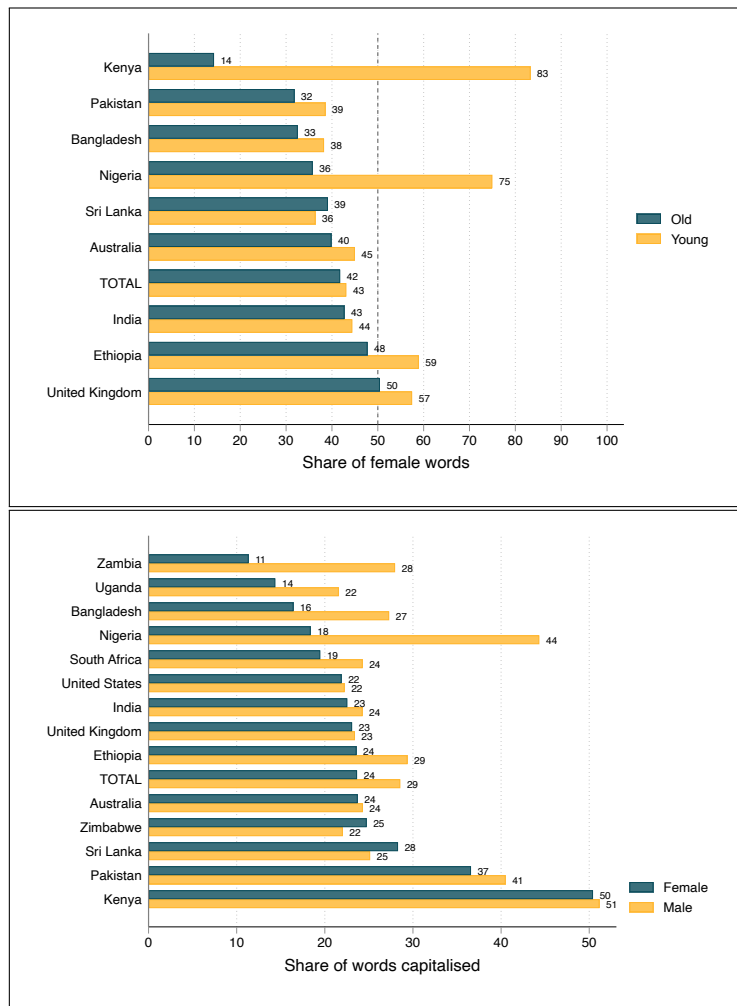
⁷Evidence shows relatively high labour market returns to study maths and science in India (Jain et al., 2022), the United States (Goodman, 2019; Levine and Zimmerman, 1995) and Denmark (Joensen and Nielsen, 2009).

Figure 2: Bias in gender representation, by country and subject



Note: These figures show the predicted mean share of gendered words that are female. This measure is first calculated for each individual book - which are then estimated as a function of country, subject, grade, and (log) book length. We exclude countries with fewer than five books in our corpus.

Figure 3: Bias in gender representation, by age and capitalisation



Note: These figures show the mean share of gendered words that are female. This measure is first calculated for each individual book. The shares by country are shown after adjusting for the subject composition of books. We exclude countries with fewer than five books in our corpus. The high-income countries are Australia, UK, and US. The lower-middle income countries with donor-funded books are Bhutan, Ethiopia, Guyana, Kenya, Lesotho, Liberia, Namibia, Nigeria, Papua New Guinea, Samoa, Sierra Leone, South Africa, South Sudan, Tonga, Zambia, and Zimbabwe.

4.2. Gender stereotypes

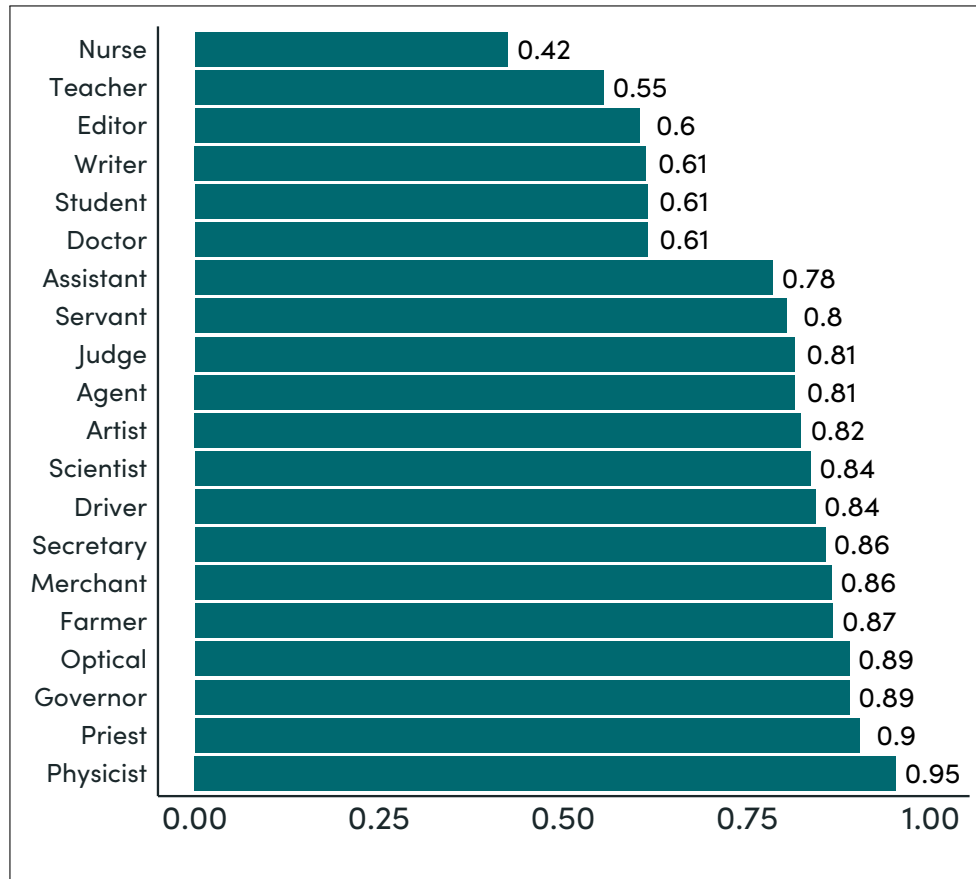
4.2.1. Word co-occurrences

In general, female words are more likely to co-occur with traditionally female occupations in our data. Because there are many more male terms in our data, most occupations have more occurrences with male terms: for example, of the 20 most commonly mentioned occupations in our corpus, all except for 'nurse' co-occur more with male terms in absolute terms (Figure 4). In relative terms, occupations that co-occur most frequently with female words include 'nurse', 'housekeeper', 'thatcher', and 'technician'

(see Figure A2). The most male-dominated occupations include 'blacksmith', 'physicist', 'astronomer', and 'mathematician'.

Occupations co-occurring with female words are relatively more likely to be managerial roles, and less likely to be manual jobs, compared to those co-occurring with male words (Figure A3). This is the case across countries with a large number of co-occurrences (Figure A5)

Figure 4: Co-occurrences with gender terms: proportion occurring with male words

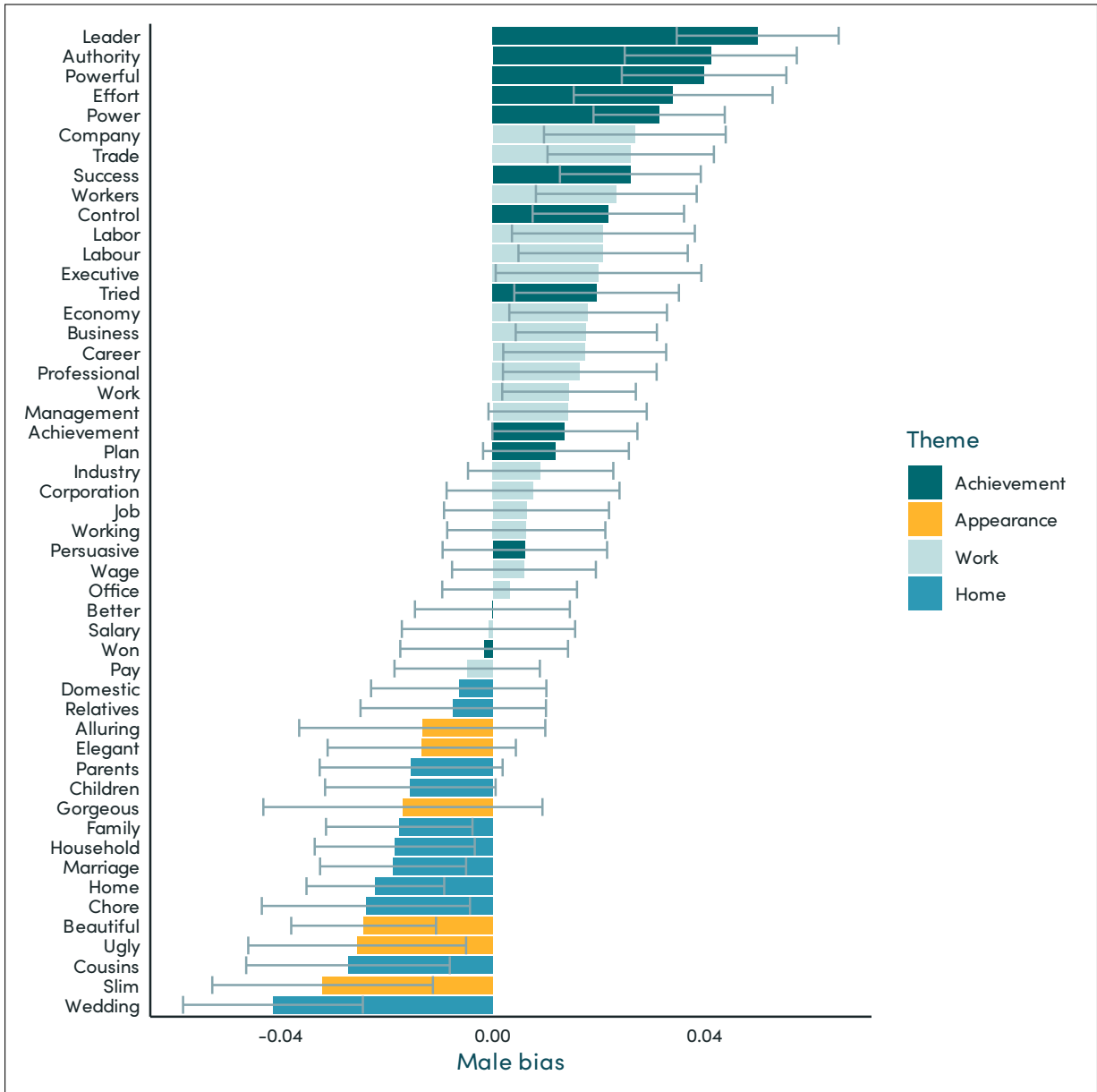


Note: This figure lists the 20 most common occupations in our full text corpus. The x-axis shows the share of all co-occurrences of the occupation and a gendered word which are male.

4.2.2. Word embeddings

Overall, we find that in line with traditional stereotypes, female words are more likely to be associated with home- or family-related words and with appearance-related words. Male words are more likely to be associated with work and achievement-related words. The words with the strongest male bias are “leader”, “authority”, and “powerful”. The words with the strongest female bias are “wedding”, “slim”, and “cousins” (Figure 5). Almost all of the individual achievement and work themed words have a stronger association with male words than female words. All of the individual appearance and home themed words have a stronger association with female words than male words.

Figure 5: Gender stereotypes



Note: This figure shows how terms relating to four themes are associated with gender terms in our embeddings. Male bias is calculated as the difference between the average cosine similarity of the theme word with the set of male gender terms, and the average similarity of the theme word with the set of female gender terms. Confidence intervals are calculated as the standard deviation for this statistic, over 50 bootstrap samples, where samples are generated by sampling all sentences in our corpus with replacement.

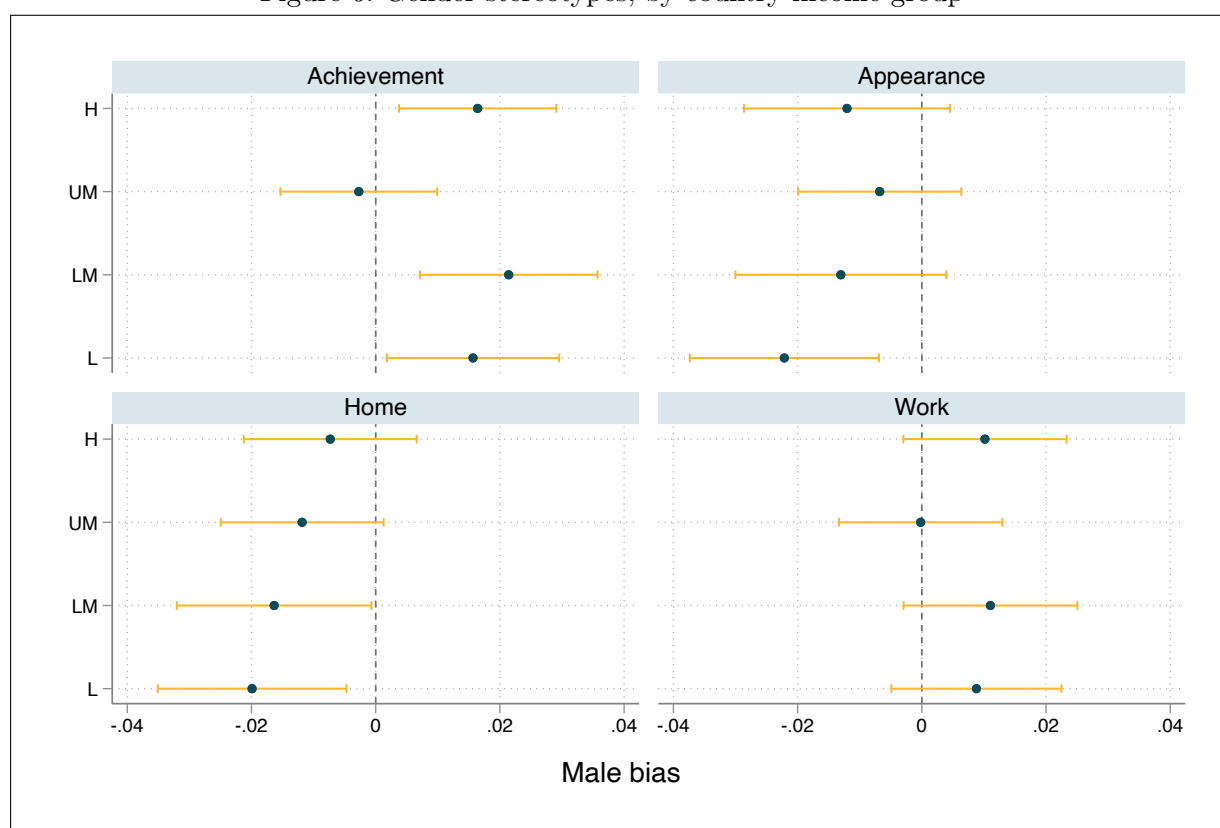
To assess heterogeneity, we consider the differences between country income groups.⁸ We pool together countries into four income groupings following the World Bank classification; low-income, lower-middle income, upper-middle income, and high-income. Overall we see similar levels of pro-male bias on average in similarity to achievement-themed and work-themed words (Figure 6). This pro-male

⁸Ash et al. (2024) recommend running word embeddings models only on a corpus with at least 1.5 million ‘tokens’ or words. Only ten countries in our sample have this many tokens. Thus we can’t look individually at all countries.

bias is though smaller in the upper-middle income countries in our sample, which are Belize, Dominica, Jamaica, Namibia, South Africa, and Tonga. All country groups have a pro-female bias in similarity with appearance- and home-themed words. This pro-female bias is largest in the books from the low-income countries in our sample (Afghanistan, Ethiopia, Rwanda, South Sudan, Liberia, Sierra Leone, and Uganda).

In Figure A6, we also show embedding biases for the ten countries in our sample with over 1.5 million tokens, which Ash et al. (2024) set as a minimum corpus size for estimating reliable embeddings. Patterns are difficult to discern, however, likely due to the still relatively small size of the corpuses.

Figure 6: Gender stereotypes, by country income group



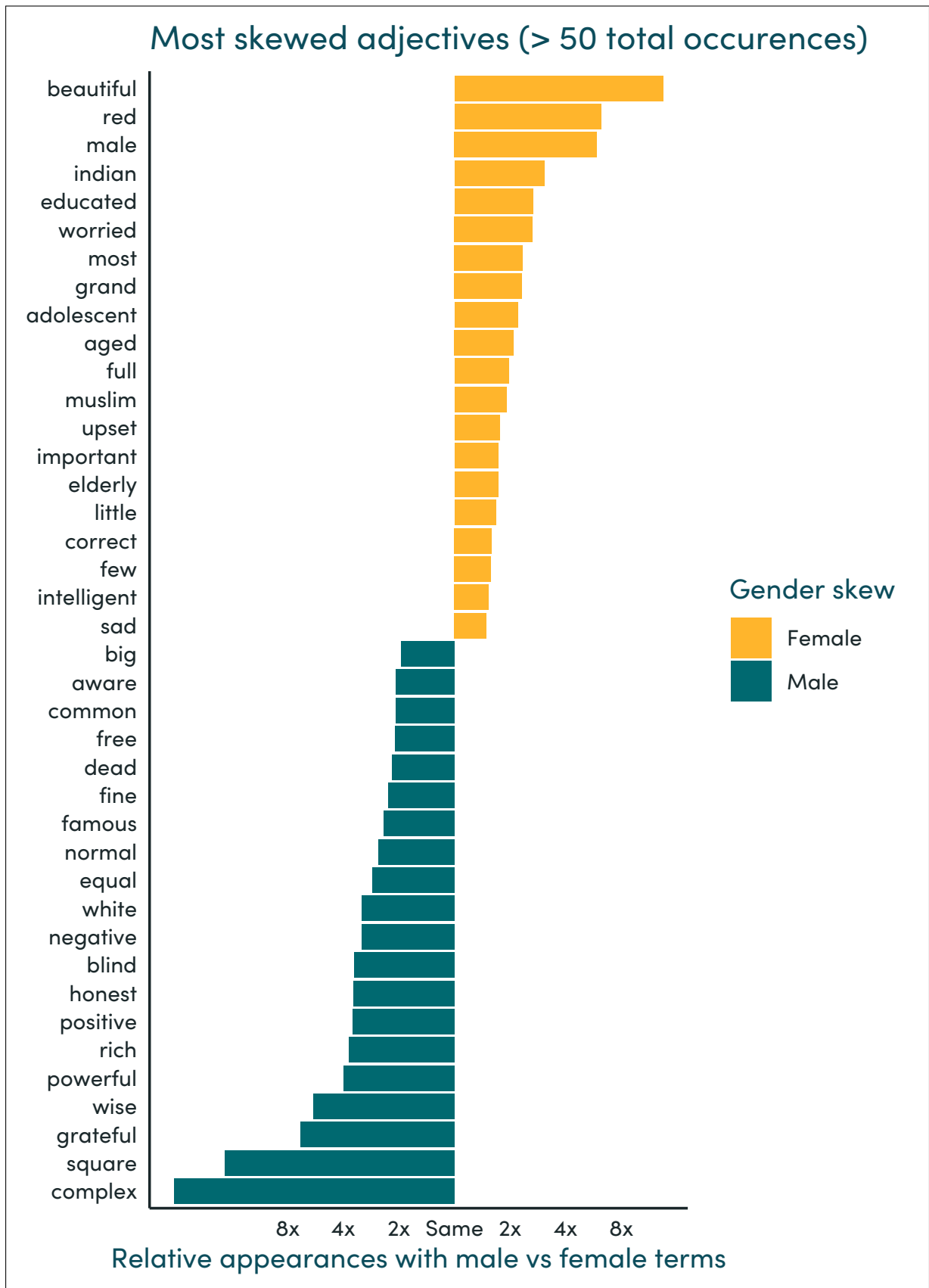
Note: Male bias is calculated as the difference between the average cosine similarity of the theme word with the set of male gender terms, and the average similarity of the theme word with the set of female gender terms. Confidence intervals are calculated as the standard deviation for this statistic, over 50 bootstrap samples, where samples are generated by sampling all sentences in our corpus with replacement, until a sample has as many sentences as the original corpus. H indicates high-income countries, which are Australia, Canada, United Kingdom, United States, Guyana, and St Kitts and Nevis. UM are upper-middle-income countries which are Belize, Dominica, Jamaica, Namibia, South Africa, and Tonga. LM are lower-middle-income countries, which are Bangladesh, India, Pakistan, Kenya, Nigeria, Zambia, Zimbabwe, Bhutan, Kiribati, Papua New Guinea, Samoa, Solomon Islands, Sri Lanka, and Lesotho. L are low-income countries, which are Afghanistan, Ethiopia, Rwanda, South Sudan, Liberia, Sierra Leone, and Uganda.

4.2.3. Part-of-speech tagging

Our part-of-speech tagging model identifies a total of 33670 sentences containing a gendered word and an adjective, 202287 sentences with a verb and a gendered word, and 164626 sentences with a noun and a gendered word across our entire corpus. In line with the over-representation of male gendered words, we also see more co-occurrences of male words in general with specific adjectives and verbs (Figure A7).

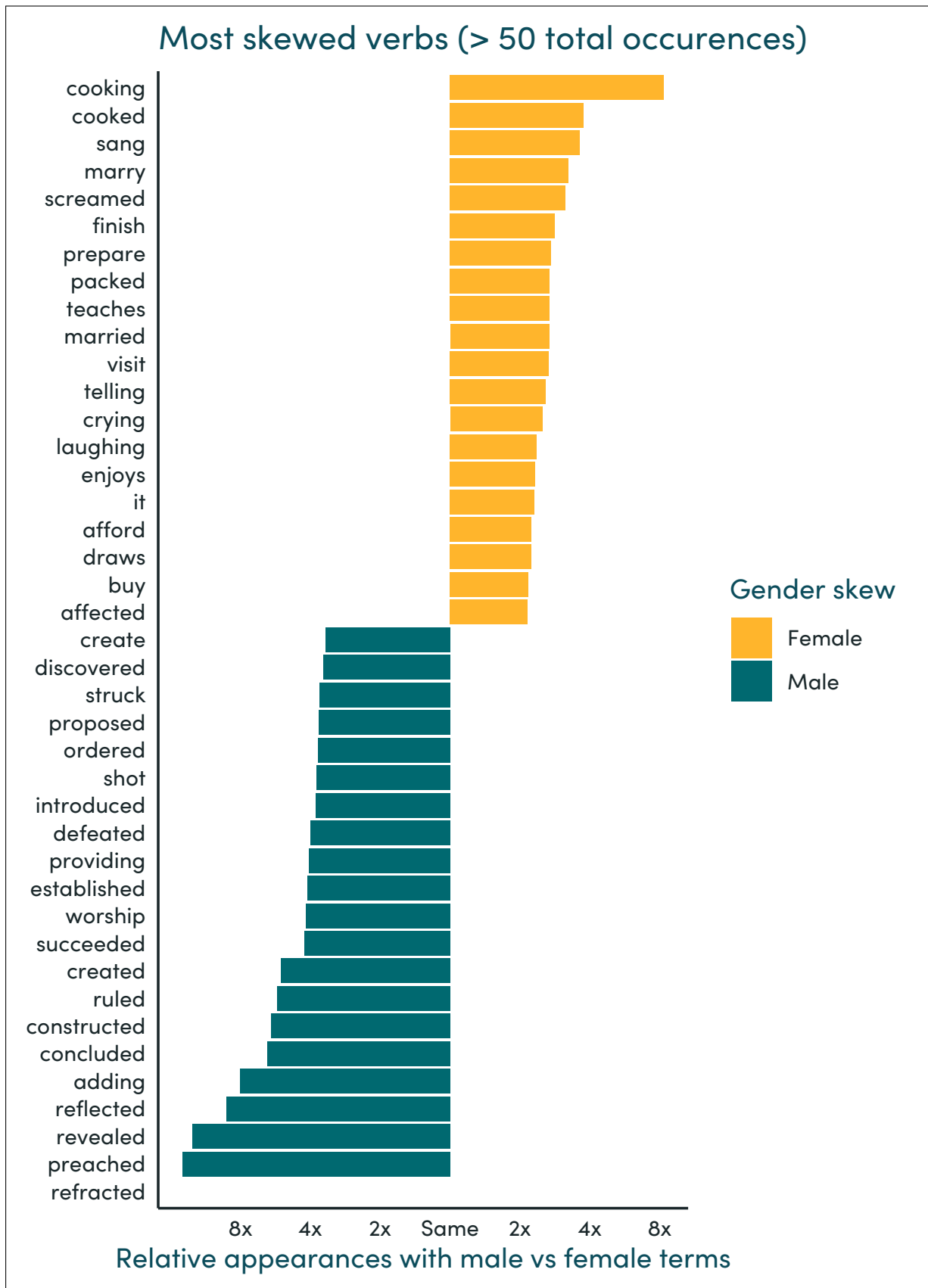
Focusing only on those words which are mentioned relatively frequently (at least 50 times in our corpus), we next estimate the ratio of mentions alongside male or female terms. The adjectives with the largest difference in likelihood in describing male over female characters include “powerful”, “rich”, “honest”, and “kind”, while verbs include “preached”, “created”, “ruled”, and “kill” (Figure 7, Figure 8). Adjectives relatively likely to describe female and not male characters include “beautiful”, “worried”, “educated” and “grand”, while verbs include “cooking”, “sang”, “marry”, and “screamed”. Due to imprecision in our analytical methods, there are also words which are likely to have been sometimes mistakenly identified as adjectives and verbs of male and female characters, such as “square”.

Figure 7: Adjectives most likely to be used for each gender



Note: This figure shows the adjectives which are most likely to be used for one gender compared to the other. After identifying adjectives used with each gender using part-of-speech-tagging, we calculate the relative frequency for each adjective, accounting for the generally higher frequency of male terms.

Figure 8: Verbs most likely to be used for each gender



Note: This figure shows the verbs which are most likely to be used for one gender compared to the other. After identifying verbs used with each gender using part-of-speech-tagging, we calculate the relative frequency for each verb, accounting for the generally higher frequency of male terms. One term, 'refracted', has been excluded from the graph, as there are no occurrences where it is used with female terms, and is therefore undefined.

After merging our annotated dataset with the Valence, Arousal, and Dominance (VAD) lexicon, we can quantify the differences between words associated with male and female characters. Overall, we find that female characters are associated with words with more positive words (valence, +0.07 standard deviations), and male characters with more active (arousal, +0.04 standard deviation difference) and dominant words (+0.08 standard deviations). These differences do not vary by country income - the gaps are as large in rich countries as in poor countries (Table 3).

Table 3: Gendered words and the valence, arousal, and dominance of co-occurring words

	(1) Valence	(2) Valence	(3) Arousal	(4) Arousal	(5) Dominance	(6) Dominance
Female	0.0720*** (0.00947)	0.123*** (0.0462)	-0.0352*** (0.00750)	0.0162 (0.0449)	-0.0772*** (0.0100)	-0.0658 (0.0503)
Log GDP		0.0108** (0.00463)		0.00280 (0.00511)		-0.00689 (0.00719)
Female X Log GDP		-0.00643 (0.00528)		-0.00629 (0.00544)		-0.00126 (0.00599)
Obs. (Sentences)	228,519	228,519	228,519	228,519	228,519	228,519
Obs. (Books)	1,096	1,096	1,096	1,096	1,096	1,096
R ²	0.001	0.001	0.000	0.000	0.001	0.001

Note: This figure presents regressions of the standardised value of the valence, arousal, and dominance of the verbs, adjectives, and nouns used in the same sentences as specified male and female terms. Data on valence, arousal, and dominance, is from Mohammad (2018). We identify verbs, nouns, and adjectives in our corpus using the UDPipe pipeline (Straka and Straková, 2017).

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

4.3. Country gender laws, norms, and textbooks

To what extent does gender bias in textbooks reflect gender-related social norms? We find a strong correlation between our measure of gender representation in textbooks and other measures of gender bias at the country level. A 1 standard deviation increase in the SIGI index, WBL index, or GEPI index is associated with a 5-6 percentage point increase in the share of gendered words in books that are female. A 1 standard deviation increase in the gender parity in secondary schooling index, or in the share of female members of parliament, is associated with a 2 percentage point increase in the share of female words in books. Increase in rates of girl marriage are associated with lower shares of female words in books. All of these correlations are estimated in regressions in which the outcome is the share of gendered words in a specific textbook that are female, estimated as a function of country-level attitudes, outcomes, and laws, whilst controlling for (log) GDP per capita, and characteristics of the textbook itself (its grade, subject, and word length) (Table 4).

Table 4: Country characteristics and bias in textbook gender representation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	share	share	share	share	share	share	share
SIGI Index (z-score)	5.655*** (0.839)						4.190 (3.514)
WBL Index (z-score)		4.965*** (1.300)					2.900 (3.755)
GEPI Index (z-score)			4.009*** (1.179)				0.264 (1.426)
Sec Ed Parity (z-score)				2.109** (0.897)			2.366** (0.906)
Girl marriage (z-score)					-2.176** (0.997)		0.812 (2.057)
Female MPs (z-score)						2.971*** (0.732)	-0.573 (1.112)
Log GDP	-1.387* (0.686)	-1.284 (0.793)	-0.755 (0.858)	0.464 (0.557)	0.229 (0.651)	0.625 (0.628)	-2.077** (0.735)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs. (Books)	919	1,191	1,141	1,191	1,121	1,112	919
Obs. (Countries)	20	29	27	29	27	26	20
R ²	0.216	0.158	0.157	0.134	0.145	0.155	0.227

Note: The outcome variable is the share of gendered words in each textbook that are female. This value can range between 0 and 100. Controls include the log of GDP per capita, as well as individual book characteristics - their subject, grade level, and word length. The SIGI Index is the OECD's Social Institutions and Gender Index. The WBL Index is the World Bank's Women, Business, and the Law Index. The GEPI Index is CGD's Girl's Education Policy Index. Data on Secondary Education Parity is from UNESCO. Data on girl marriage and female MPs is from the OECD's Gender, Institutions, and Development database.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

5. Discussion: How much does bias in textbooks really matter?

Written materials, including textbooks, curricula, and examinations, play a major role in schooling. Thus, how textbooks show girls and boys, women and men, in various different roles, and the portrayal of males and females might influence how children and adolescents view gender roles in society. But how important are they?

The main theory of gender development posits that environmental factors interact with three major social modes of influence; (i) modelling by people in one's immediate environment and through mass media, (ii) through social reactions to one's own behaviour, and (iii) through direct tuition. These factors are all expected to interact in complex ways, but with the largest influence coming from modelling rather than direct tuition (Bussey and Bandura, 1999). We also know that children's beliefs about gender stereotypes start young. One study showed that boys and girls as young as six years old thought that boys were more likely to be 'brilliant' than girls (Bian et al., 2017).

Evidence on the causal impact of bias in textbooks on children is limited. One exception is a study from Bangladesh that showed that providing children with gender sensitive storybooks made them more likely to agree that boys and girls should do the same things when they grow up, both do chores, and both play football (Leighton and Mucyo, 2019). Small-scale lab experiments in the United States have shown that children express a preference for same-gender characters in books, and their behaviour is more influenced by the behaviour of same-gender characters in storybooks (McArthur and Eisen, 1976; Ashton, 1983). Children taught with gender-stereotyped books are more likely to express gender-stereotyped views about appropriate activities, personality characteristics, and career roles (Karniol and Gal-Disegni, 2009; Peterson and Lach, 1990). Experimental evidence also shows that gender stereotypes can be reduced through exposure to counter-examples (Parish and Bryant, 1976; Nhundu, 2007).

A growing literature in economics documents the importance of role models and aspirations in determining schooling and employment outcomes, which could plausibly be provided by fictional characters as much as in-person interactions (Serra, 2022), as shown for instance by Ochman (1996). Experimental evidence from Australia found that showing books with same-sex role models to grade three children increased self-esteem, for both boys and girls (Ochman, 1996). This notion is also supported by survey data from Kenya showing that children were more likely to empathise with characters in books of their own gender (Mburu and Nyagah, 2012).

Outside of the classroom, gendered language in school examinations has been shown to lead to worse performance for girls (Cohen et al., 2023). Countries with gendered national languages have worse employment and education outcomes for women (Jakiela and Ozier, 2020).

6. Conclusions

In this paper we gather a new corpus of 1255 textbooks from 34 countries, and document under-representation of female characters and gendered stereotypes. Textbooks from low and lower-middle income countries show a greater degree of bias than upper middle and high-income countries. Lower female representation in books correlates with other measures of gender equality.

There are several limitations to our analysis. First, we focus only on text and not images. Advances in computer image recognition allow for the identification of different characters in pictures, though this is beyond the scope of our paper. Second, we focus exclusively on English language books, though our approach is in principle compatible with other languages. Third, while we manually validate the automated processes we use where possible, several of our analytical methods entail some imprecision in measuring bias. For example, some 'names' identified may not be names in context, and the co-

occurrence of a gender term with an occupation term will not always indicate they are connected. Fourth, we are not able to distinguish sections of the text which are read by students to sections which are not, such as information about the book publisher.

Our analysis highlights which countries and subjects have particularly low female representation. This presents a new front in the girl's education policy agenda. Several countries in our database have achieved gender parity in school enrolment, but still see significant under-representation of women and girls in school materials. Paying particular attention to gender representation could be prioritised in future revisions of these books. A constraint to such reform is the likelihood that reform is most needed in countries in which it is least likely due to regressive gender norms amongst adults. This is in part addressed in the countries in which government textbooks are partly funded by donors.

As well as highlighting the scope for revision of existing books, our analysis can also be easily replicated with new books, and so could be used as a proactive check for new drafts.

Our analysis was limited by the public availability of books, even where they were publicly funded by governments with policies in place mandating open access to materials. Donors and national governments should ensure that all publicly funded school books are made easily available online in digital copy. Digital books are public goods which should be freely available. Book availability to households is particularly important in the context of increasingly frequent school closures due to climate-related shocks, as well as shocks related to conflict and disease outbreaks. The opening up of schoolbooks to researchers is just an added benefit.

Future research could easily expand our analysis to more datasets. An advantage of the quantitative approach is that our analysis is easily scalable, and can be extended to new corpuses of digitized books as they become available. New Analysis of image data would also be valuable. As our analysis is largely descriptive, new experimental research would be useful in documenting the impact of gender bias in books on student attitudes. Several authors have argued that replacing textbooks with more progressive ones may not be enough to change attitudes if teachers do not engage with them (Kostas, 2023; Namatende-Sakwa, 2021).

References

- Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz**, “What We Teach About Race and Gender: Representation in Images and Text of Children’s Books*,” *The Quarterly Journal of Economics*, November 2023, 138 (4), 2225–2285.
- , **Patricia Chiril, Callista Christ, Anjali Das, Alex Eble, Emileigh Harrison, and Hakizumwami Birali Runesha**, “Tales and Tropes: Gender Roles from Word Embeddings in a Century of Children’s Books,” Technical Report 2022.
- Asadullah, M. Niaz, Kazi Md Mukitul Islam, and Zaki Wahhaj**, “Gender Bias in Bangladeshi School Textbooks: Not Just a Matter of Politics or Growing Influence of Islamists,” *The Review of Faith & International Affairs*, April 2018, 16 (2), 84–89.
- Ash, Elliott, Daniel L Chen, and Arianna Ornaghi**, “Gender Attitudes in the Judiciary: Evidence from U.S. Circuit Courts,” *American Economic Journal: Applied Economics*, 2024.
- Ashton, Eleanor**, “Measures of play behavior: The influence of sex-role stereotyped children’s books,” *Sex Roles*, January 1983, 9 (1), 43–47.
- Awich, Morris**, “SACMEQ IV International Report,” Technical Report, The Southern and Eastern Africa Consortium for Monitoring Educational Quality 2021.
- Bachore, Mebratu and Tesfaye Semela**, “Is gender-bias in textbooks spilled over from schools to universities? Evidence from Ethiopian university English language course textbooks,” *Cogent Education*, December 2022, 9 (1), 2152616. Publisher: Cogent OA eprint: <https://doi.org/10.1080/2331186X.2022.2152616>.
- Bahman, Masoumeh**, “The Sexist Occupational Representation in English Textbooks of Iranian High schools | Issues in Language Studies,” December 2020.
- Barton, Amanda and Lydia Namatende Sakwa**, “The representation of gender in English textbooks in Uganda,” *Pedagogy, Culture & Society*, July 2012, 20 (2), 173–190.
- Bhattarai, Swechhya**, “Gender Mainstreaming on Nepal’s Pedagogy from 1995 - 2020,” Background paper prepared for the Global Education Monitoring Report Gender Report 2020.
- Bian, Lin, Sarah-Jane Leslie, and Andrei Cimpian**, “Gender stereotypes about intellectual ability emerge early and influence children’s interests,” *Science*, January 2017, 355 (6323), 389–391. Publisher: American Association for the Advancement of Science.
- Bies, Ann, Mott, Justin, Warner, Colin, and Kulick, Seth**, “English Web Treebank,” August 2012. Artwork Size: 105472 KB Pages: 105472 KB.

- Bietenbeck, Jan, Natalie Irmert, and Mohammad H. Sepahvand**, “Teacher quality and cross-country differences in learning in francophone Sub-Saharan Africa,” *Economics of Education Review*, October 2023, *96*, 102437.
- Blumberg, Rae Lesser**, “The Invisible Obstacle to Educational Equality: Gender Bias in Textbooks,” *Prospects: Quarterly Review of Comparative Education*, September 2008, *38* (3), 345–361. Publisher: Springer ERIC Number: EJ849088.
- Bolukbasi, Tolga, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai**, “Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings,” in “Advances in Neural Information Processing Systems,” Vol. 29 Curran Associates, Inc. 2016.
- Boutyline, Andrei, Alina Arseniev-Koehler, and Devin Cornell**, “School, Studying, and Smarts: Gender Stereotypes and Education Across 80 Years of American Print Media, 1930-2009,” *Social Forces*, 2023.
- Brugailles, Carole and Sylvie Cromer**, *Analysing gender representations in school textbooks*, english edition ed., Paris: Population and Development Centre-Unit 196 Université Paris Descartes-INED-IRD, 2009.
- Buckner, Elizabeth and Susan Garnett Russell**, “Portraying the Global: Cross-national Trends in Textbooks’ Portrayal of Globalization and Global Citizenship1,” *International Studies Quarterly*, December 2013, *57* (4), 738–750.
- Bussey, Kay and Albert Bandura**, “Social cognitive theory of gender development and differentiation.,” *Psychological Review*, 1999, *106* (4), 676–713.
- Cantoni, Davide and Noam Yuchtman**, “The political economy of educational content and development: Lessons from history,” *Journal of Development Economics*, September 2013, *104*, 233–244. 20 citations (Crossref) [2022-11-07].
- , **Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang**, “Curriculum and Ideology,” *Journal of Political Economy*, April 2017, *125* (2), 338–392.
- Chen, Wei-Lin, Ming-Jen Lin, and Tzu-Ting Yang**, “Curriculum and national identity: Evidence from the 1997 curriculum reform in Taiwan,” *Journal of Development Economics*, March 2023, p. 103078.
- Cohen, Alma, Tzur Karelitz, Tamar Kricheli-Katz, Sephi Pumpian, and Tali Regev**, “Gender-Neutral Language and Gender Disparities,” April 2023.
- Crawfurd, Lee and Susannah Hares**, “Introducing the Girl’s Education Policy Index,” 2020.

- de Marneffe, Marie-Catherine, Christopher D. Manning, Joakim Nivre, and Daniel Zeman,** “Universal Dependencies,” *Computational Linguistics*, July 2021, *47* (2), 255–308.
- Dhar, Diva, Tarun Jain, and Seema Jayachandran,** “Reshaping Adolescents’ Gender Attitudes: Evidence from a School-Based Experiment in India,” *American Economic Review*, March 2022, *112* (3), 899–927.
- Fredriksen, Birger and Sukhdeep Brar,** “Getting Textbooks to Every Child in Sub-Saharan Africa,” Technical Report, World Bank 2015.
- Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou,** “Word embeddings quantify 100 years of gender and ethnic stereotypes,” *Proceedings of the National Academy of Sciences*, April 2018, *115* (16), E3635–E3644. Publisher: Proceedings of the National Academy of Sciences.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin,** “Many Children Left Behind? Textbooks and Test Scores in Kenya,” *American Economic Journal: Applied Economics*, January 2009, *1* (1), 112–135.
- Goodman, Joshua,** “The Labor of Division: Returns to Compulsory High School Math Coursework,” *Journal of Labor Economics*, October 2019, *37* (4), 1141–1182. Publisher: The University of Chicago Press.
- Hegghammer, Thomas,** “OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment,” *Journal of Computational Social Science*, May 2022, *5* (1), 861–882.
- Islam, Kazi Md Mukitul and M. Niaz Asadullah,** “Gender stereotypes and education: A comparative content analysis of Malaysian, Indonesian, Pakistani and Bangladeshi school textbooks,” *PLOS ONE*, January 2018, *13* (1), e0190807. Publisher: Public Library of Science.
- Jain, Tarun, Abhiroop Mukhopadhyay, Nishith Prakash, and Raghav Rakesh,** “Science education and labor market outcomes in a developing economy,” *Economic Inquiry*, April 2022, *60* (2), 741–763.
- Jakiela, Pamela and Owen W. Ozier,** “Gendered Language,” April 2020.
- Jimenez, Jeremy David, Julia Lerch, and Patricia Bromley,** “Education for global citizenship and sustainable development in social science textbooks,” *European Journal of Education*, 2017, *52* (4), 460–476. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ejed.12240>.
- Joensen, Juanna Schrøter and Helena Skyt Nielsen,** “Is there a Causal Effect of High School Math on Labor Market Outcomes?,” *Journal of Human Resources*, January 2009, *44* (1), 171–198. Publisher: University of Wisconsin Press Section: Articles.

- Karama, Muneer J**, “Gender Bias in School Mathematics Textbooks from Grade 1 to 12 in Palestine,” *Journal of International Women’s Studies*, 2020, *21* (1).
- Karniol, Rachel and Michal Gal-Disegni**, “The Impact of Gender-Fair versus Gender-Stereotyped Basal Readers on 1st-Grade Children’s Gender Stereotypes: A Natural Experiment,” *Journal of Research in Childhood Education*, June 2009, *23* (4), 411–420. Publisher: Routledge .eprint: <https://doi.org/10.1080/02568540909594670>.
- Kincheloe, Joe L.**, *Understanding the New Right and Its Impact on Education*, Phi Delta Kappa Educational Foundation, 1983. Google-Books-ID: M4MkAQAAAMAAJ.
- Kostas, Marios**, “Textbooks, students and teachers talk around gender: A new materialist approach to children’s agency,” *Teaching and Teacher Education*, April 2023, *125*, 104052.
- Kuecken, Maria and Marie-Anne Valfort**, “When do textbooks matter for achievement? Evidence from African primary schools,” *Economics Letters*, June 2013, *119* (3), 311–315.
- Kumar, Vaibhav, Tenzin Singhay Bhotia, Vaibhav Kumar, and Tanmoy Chakraborty**, “Identifying and Mitigating Gender Bias in Hyperbolic Word Embeddings,” September 2021. arXiv:2109.13767 [cs].
- Lee, Boyoon**, “Educational Content, Exclusive National Identity, and Anti-Immigrant Attitudes,” *The Journal of Politics*, March 2023. Publisher: The University of Chicago Press.
- Lee, Jackie F.K. and Peter Collins**, “Construction of gender: a comparison of Australian and Hong Kong English language textbooks,” *Journal of Gender Studies*, June 2010, *19* (2), 121–137.
- Leighton, Margaret and Kirsten Mucyo**, “Bangladesh Children’s Book Initiative Endline Report,” Technical Report, Save the Children 2019.
- Levine, Phillip B. and David J. Zimmerman**, “The Benefit of Additional High-School Math and Science Classes for Young Men and Women,” *Journal of Business & Economic Statistics*, 1995, *13* (2), 137–149. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].
- Lewis, Molly, Matt Cooper Borkenhagen, Ellen Converse, Gary Lupyan, and Mark S. Seidenberg**, “What Might Books Be Teaching Young Children About Gender?,” *Psychological Science*, January 2022, *33* (1), 33–47. Publisher: SAGE Publications Inc.
- Lucy, Li, Dorottya Demszky, Patricia Bromley, and Dan Jurafsky**, “Content Analysis of Textbooks via Natural Language Processing: Findings on Gender, Race, and Ethnicity in Texas U.S. History Textbooks,” *AERA Open*, July 2020, *6* (3), 2332858420940312.

- Martin, Michael, Chad A. Minnich, Ina V. S. Mullis, Gabrielle M. Stanco, Alka Arora, Victoria A.S Centurino, and Courtney E. Castle**, *Timss 2011 encyclopedia: education policy and curriculum in mathematics and science*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, 2012.
- Maruyama, Takao and Takashi Kurosaki**, “Developing Textbooks to Improve Math Learning in Primary Education: Empirical Evidence from El Salvador,” *Economic Development and Cultural Change*, July 2022, pp. 000–000. Publisher: The University of Chicago Press.
- Mburu, David N. P. and Grace Nyagah**, “Effects of Gender Role Portrayal in Textbooks in Kenyan Primary Schools, on Pupils Academic Aspirations,” *Problems of Education in the 21st Century*, November 2012, 47 (1), 100–109.
- McArthur, Leslie Z. and Susan V. Eisen**, “Achievements of male and female storybook characters as determinants of achievement behavior by boys and girls,” *Journal of Personality and Social Psychology*, 1976, 33 (4), 467–473. Place: US Publisher: American Psychological Association.
- Mehmood, Sultan, Shaheen Naseer, and Daniel L Chen**, “Transmitting Rights: Effective Cooperation, Inter-Gender Contact, and Student Achievement,” Technical Report 2022.
- Menendez, David Arroyo**, “Damegender: Towards an International and Free Dataset about Name, Gender and Frequency,” 2022.
- Miguel, Edward**, “Tribe or nation? Nation building and public goods in Kenya versus Tanzania,” *World Politics*, 2004, 56 (03), 328–362.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean**, “Distributed Representations of Words and Phrases and their Compositionality,” October 2013. arXiv:1310.4546 [cs, stat].
- Milligan, Lizzi O., Leon Tikly, Timothy Williams, Jean-Marie Vianney, and Alphonse Uworwabayeho**, “Textbook availability and use in Rwandan basic education: A mixed-methods study,” *International Journal of Educational Development*, May 2017, 54, 1–7.
- Mirza, Munawar**, “Gender analysis of school curriculum and text books; 2004,” Technical Report, UNESCO, Islamabad 2004.
- Mohammad, Saif**, “Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words,” in “Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)” Association for Computational Linguistics Melbourne, Australia 2018, pp. 174–184.

- Mustapha, Abolaji S. and Sara Mills**, *Gender Representation in Learning Materials: International Perspectives*, Routledge, August 2015. Google-Books-ID: pgRXCgAAQBAJ.
- Mustapha, Abolaji Samuel**, “Gender biased representation in learning materials in Nigeria educational system,” *Multidisciplinary Journal of Gender Studies*, June 2014, 3 (2), 368–392. Number: 2.
- Namatende-Sakwa, Lydia**, “Are Progressive Texts Necessarily Disruptive? Investigating Teacher Engagement with Gendered Textbooks in Ugandan Classrooms,” *Teachers College Record*, January 2021, 123 (1), 1–26. Publisher: SAGE Publications.
- Nhundu, Tichatonga J.**, “Mitigating Gender-typed Occupational Preferences of Zimbabwean Primary School Children: The Use of Biographical Sketches and Portrayals of Female Role Models,” *Sex Roles*, May 2007, 56 (9), 639–649.
- Ochman, Jan M.**, “The effects of nongender-role stereotyped, same-sex role models in storybooks on the self-esteem of children in grade three,” *Sex Roles*, December 1996, 35 (11), 711–735.
- Parish, Thomas S. and William T. Bryant**, “Reversing the Effects of Sexism in Elementary School Girls Through the Use of Counterconditioning Procedures,” *Classroom Interaction Newsletter*, 1976, 11 (2), 7–15. Publisher: Journal of Classroom Interaction.
- Peterson, Sharyl Bender and Mary Alyce Lach**, “Gender Stereotypes in Children’s Books: their prevalence and influence on cognitive and affective development,” *Gender and Education*, January 1990, 2 (2), 185–197. Publisher: Routledge eprint: <https://doi.org/10.1080/0954025900020204>.
- Piper, Benjamin, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, and Simon J. King**, “Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers’ guides,” *World Development*, June 2018, 106, 324–336.
- Read, Tony**, “Where Have All the Textbooks Gone?: Toward Sustainable Provision of Teaching and Learning Materials in Sub-Saharan Africa,” Technical Report, World Bank, Washington, DC 2015.
- Rodriguez, Pedro L. and Arthur Spirling**, “Word Embeddings: What Works, What Doesn’t, and How to Tell the Difference for Applied Research,” *The Journal of Politics*, November 2021, pp. 000–000. 13 citations (Crossref) [2022-11-07] ZSCC: 0000005.
- Russell, S. Garnett, Lydia Namatende-Sakwa, and Sarah Lewinger**, “Gender Representation in East African History Textbooks: The Cases of Rwanda, Kenya, and Uganda,” *Feminist Formations*, 2021, 33 (2), 233–261.

- Sadker, David M. and Karen Zittleman**, “Practical Strategies for Detecting and Correcting Gender Bias in Your Classroom,” in “Gender in the classroom: Foundations, skills, methods, and strategies across the curriculum,” Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2007, pp. 259–275.
- Serra, Danila**, “Role Models in Developing Countries,” In Preparation for the Handbook of Experimental Development Economics 2022.
- Sheridan, John**, “occupations,” February 2015. original-date: 2015-02-05T20:51:11Z.
- Silveira, Natalia, Timothy Dozat, Christopher Manning, Sebastian Schuster, Ethan Chi, John Bauer, Miriam Connor, Marie-Catherine de Marneffe, Nathan Schneider, Sam Bowman, Hanzhi Zhu, Daniel Galbraith, and John Bauer**, “Universal Dependencies English Web Treebank,” 2023.
- Singla, Shikhar and Mayukh Mukhopadhyay**, “Gender Norms Do Not Persist But Converge Across Time,” May 2022.
- Skinner, Nadine and Patricia Bromley**, “Individual and Collective Social Justice Education: Comparing Emphases on Human Rights and Social Movements in Textbooks Worldwide,” *Comparative Education Review*, November 2019, 63 (4), 502–528.
- Stará, Jana, Martin Chvál, and Karel Starý**, “The Role of Textbooks in Primary Education,” *e-Pedagogium*, December 2017, 17 (4), 60–69.
- Stern, Jonathan, Matthew Jukes, Joseph DeStefano, Jessica Mejia, Bidemi Carrol, Rachel Jordan, Cosmos Gatuyu, Tabitha Nduku, Christine Van Keuren, Maitri Punjabi, and Fatima Tufail**, “Learning at Scale: Final Report,” Technical Report, RTI International October 2023.
- Straka, Milan and Jana Straková**, “Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe,” in “Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies” Association for Computational Linguistics Vancouver, Canada 2017, pp. 88–99.
- Ward, Michael**, “PISA for Development: Results in Focus,” Technical Report, OECD, Paris December 2018.
- Witt, Susan D.**, “The Influence of School Texts on Children’s Gender Role Socialization,” *Curriculum and Teaching*, January 2001, 16 (1), 25–43.
- World Bank**, “Education Service Delivery Indicators,” 2023.

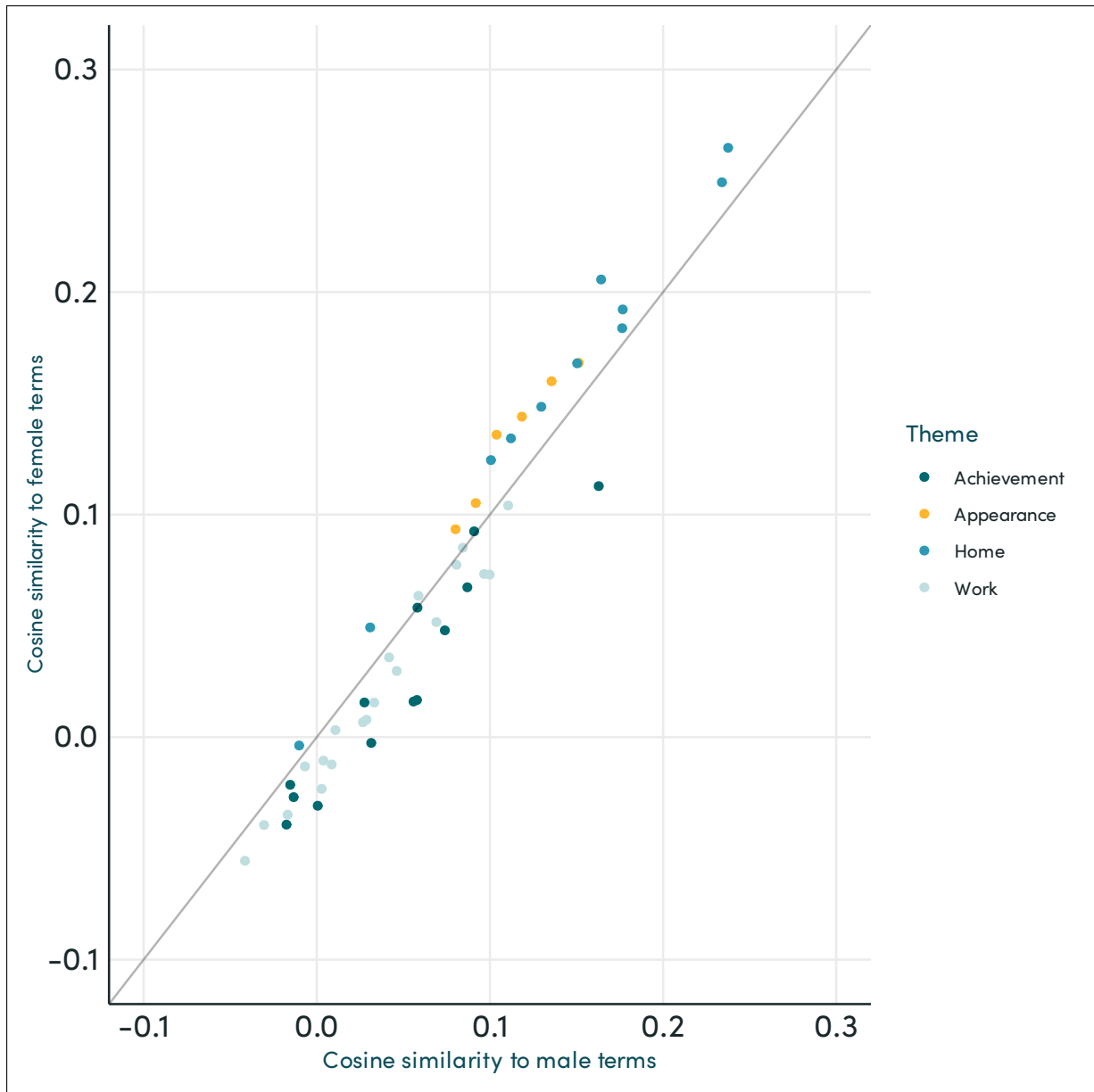
Ye, Wangbei, “The Return of “One Guideline, One Textbook” Policy: Moral Education Textbook and Teacher Interaction in China,” *Comparative Education Review*, January 2022, pp. 000–000.

İncikabi and Ulusoy

İncikabi, Lütfi and Fadime Ulusoy, “Gender bias and stereotypes in Australian, Singaporean and Turkish mathematics textbooks,” *Turkish Journal of Education*, October 2019, 8 (4), 298–317.

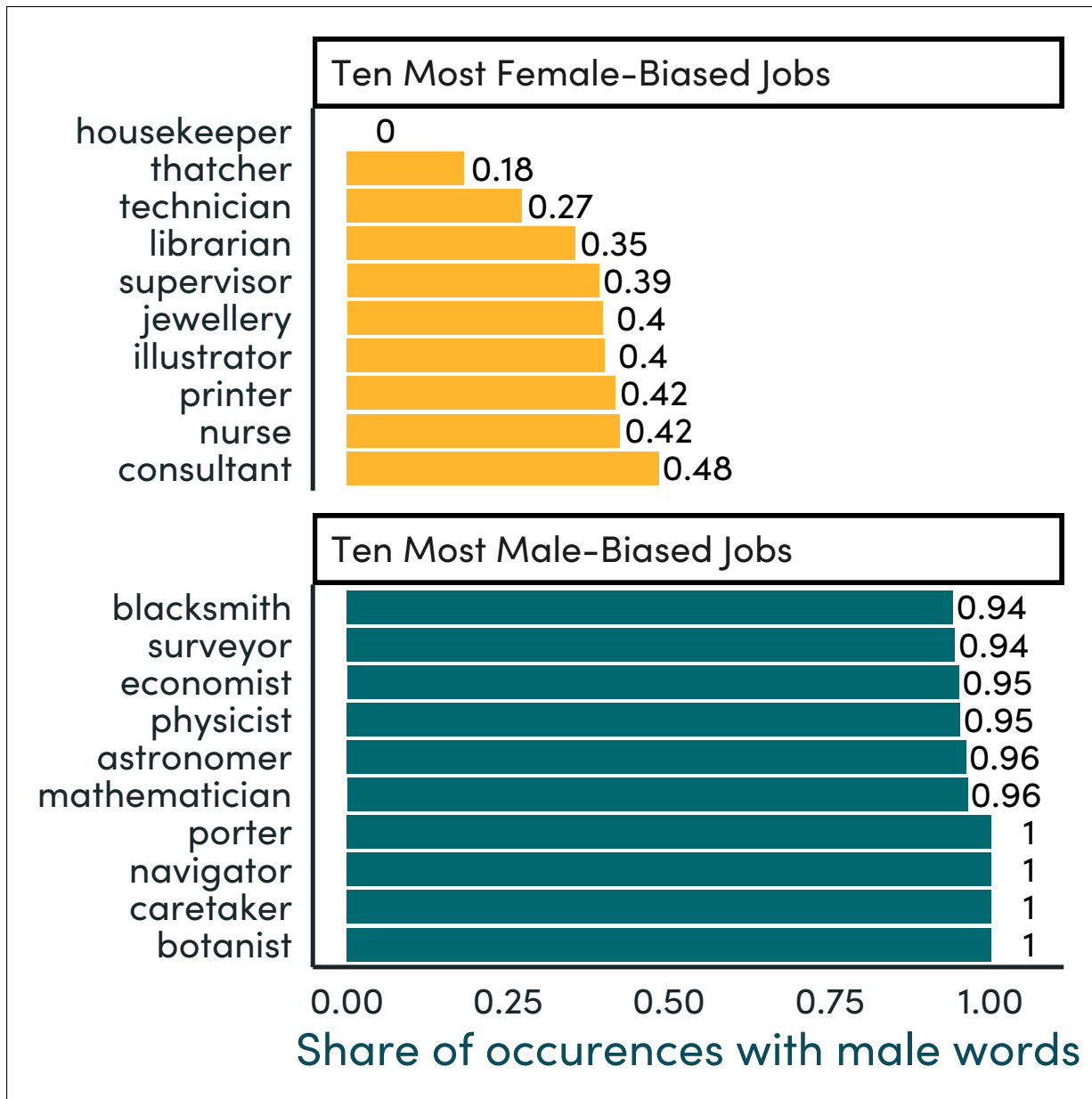
A. Appendix: Figures

Figure A1: Gendered word associations



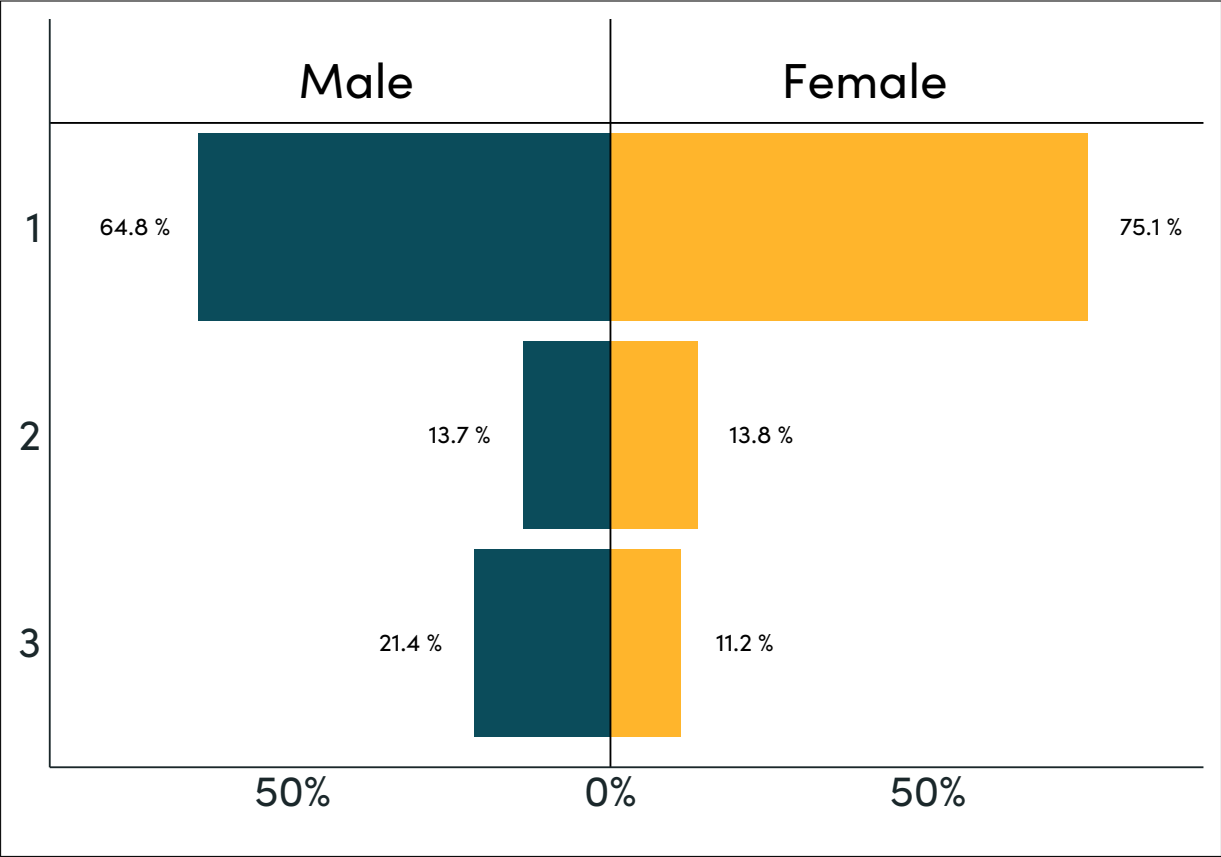
Note: Words above the 45° line are discussed more frequently in relation to women and girls, words below the line more frequently in relation to men and boys.

Figure A2: Gender-biased occupations



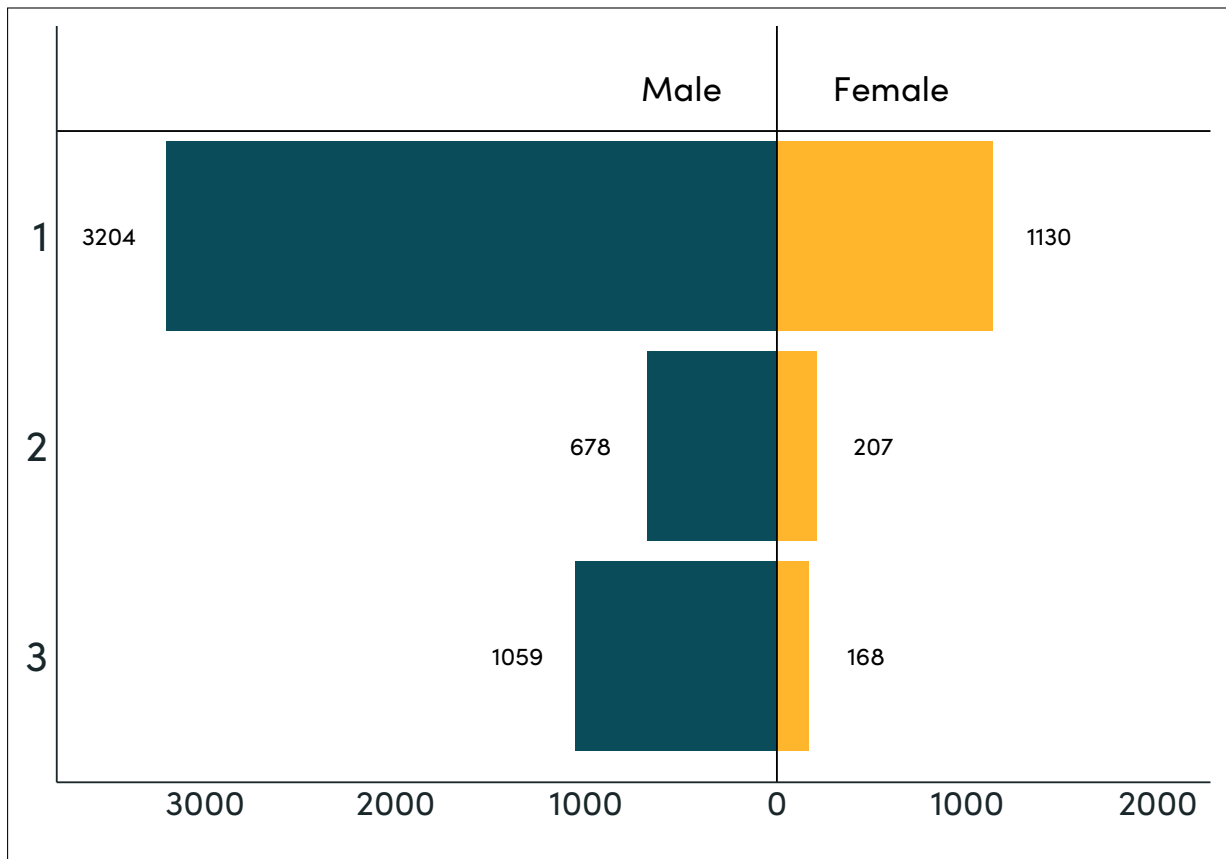
Note: This figure shows common occupation terms - classified as those with more than 10 occurrences across the corpus - which were most biased towards either gender, in the sense of co-occurring (within sentences) more frequently with terms of one than the other.

Figure A3: Occupation terms occurring with gendered terms, by ISCO code



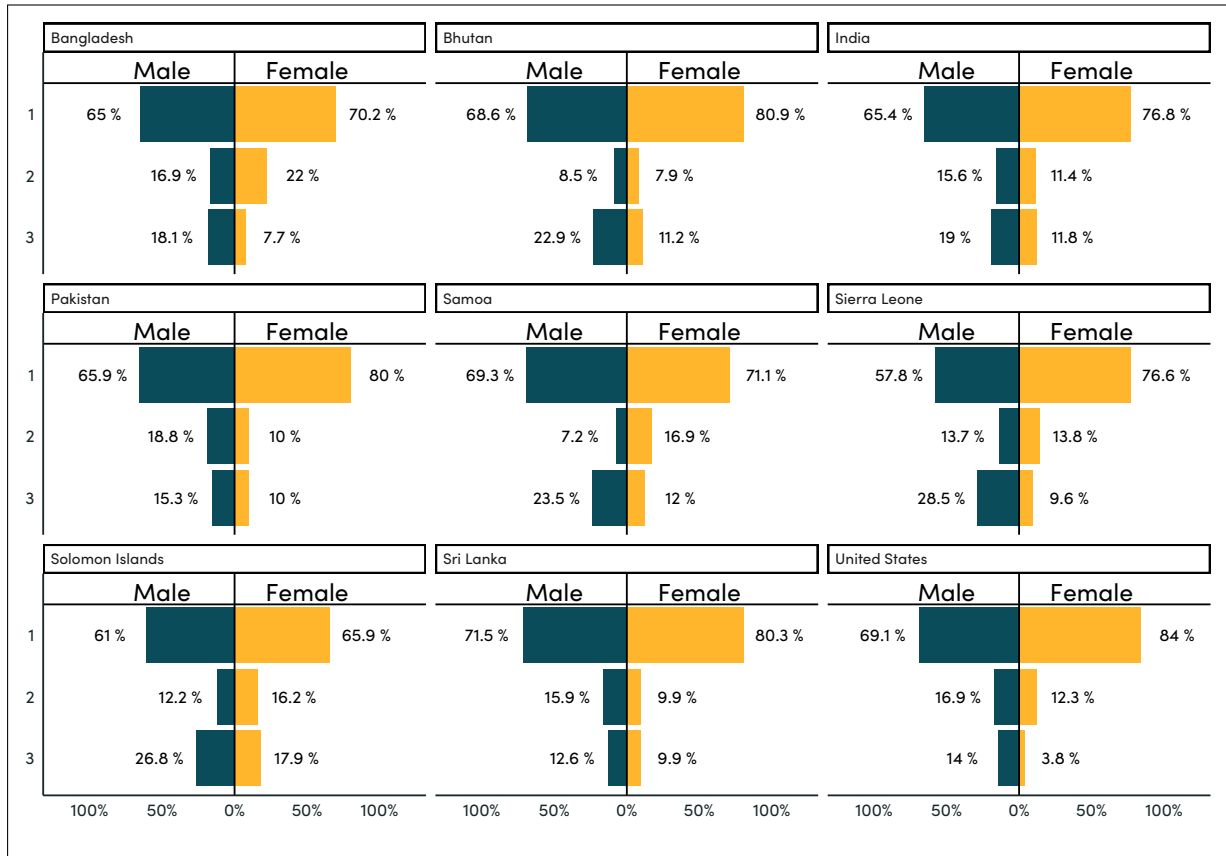
Note: This figure shows the distribution of occupation categories across co-occurrences with gendered words. That is, for all co-occurrences between a male gendered word and an occupation, 64.8 percent of those co-occurrences were with managerial or professional occupations (category 1-2). 13.7 percent with service (category 3) occupations, and 21.4 percent with manual (category 4) occupations.

Figure A4: Occupation terms occurring with gendered terms, by ISCO code, absolute counts



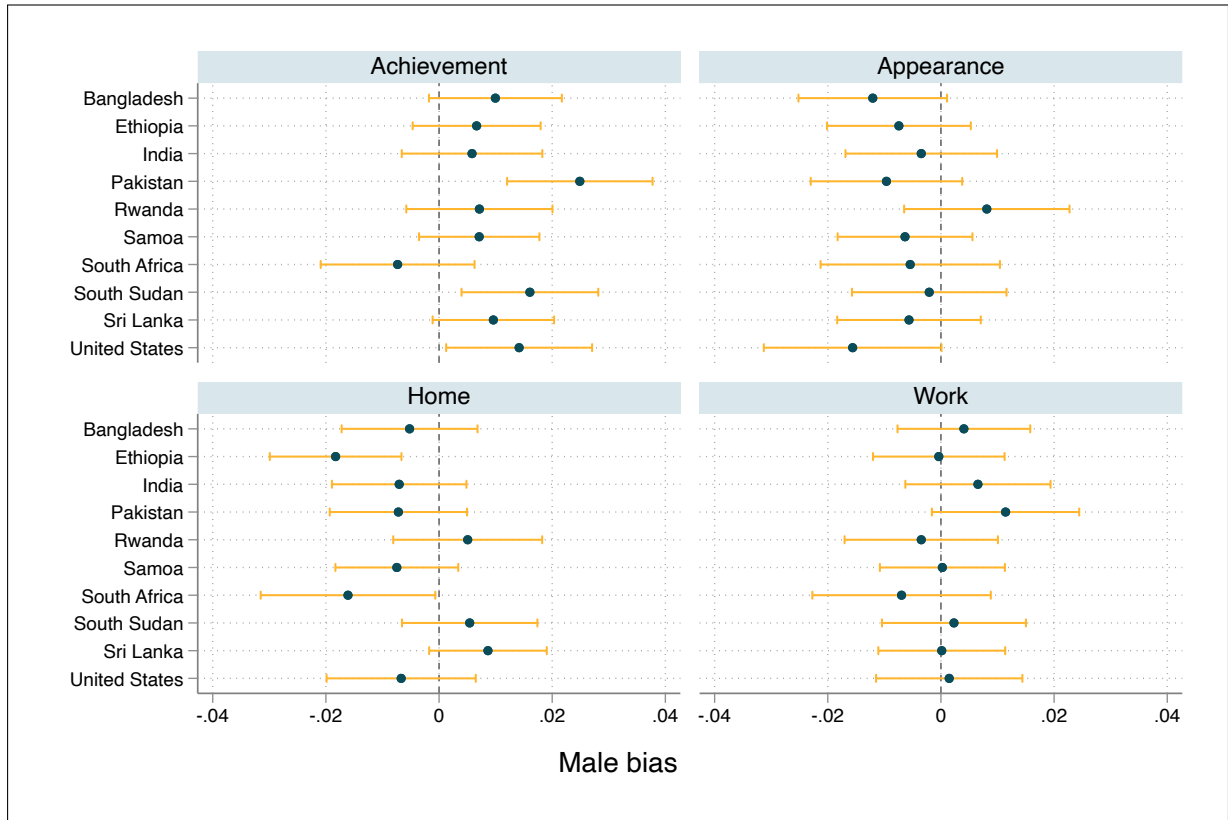
Note: This figure shows the distribution of occupation categories across co-occurrences with gendered words. In this case, absolute values are shown. For example, there were 3204 co-occurrences between male gendered words with a managerial (category 1) occupation.

Figure A5: Occupation terms occurring with gendered terms, by ISCO code and country



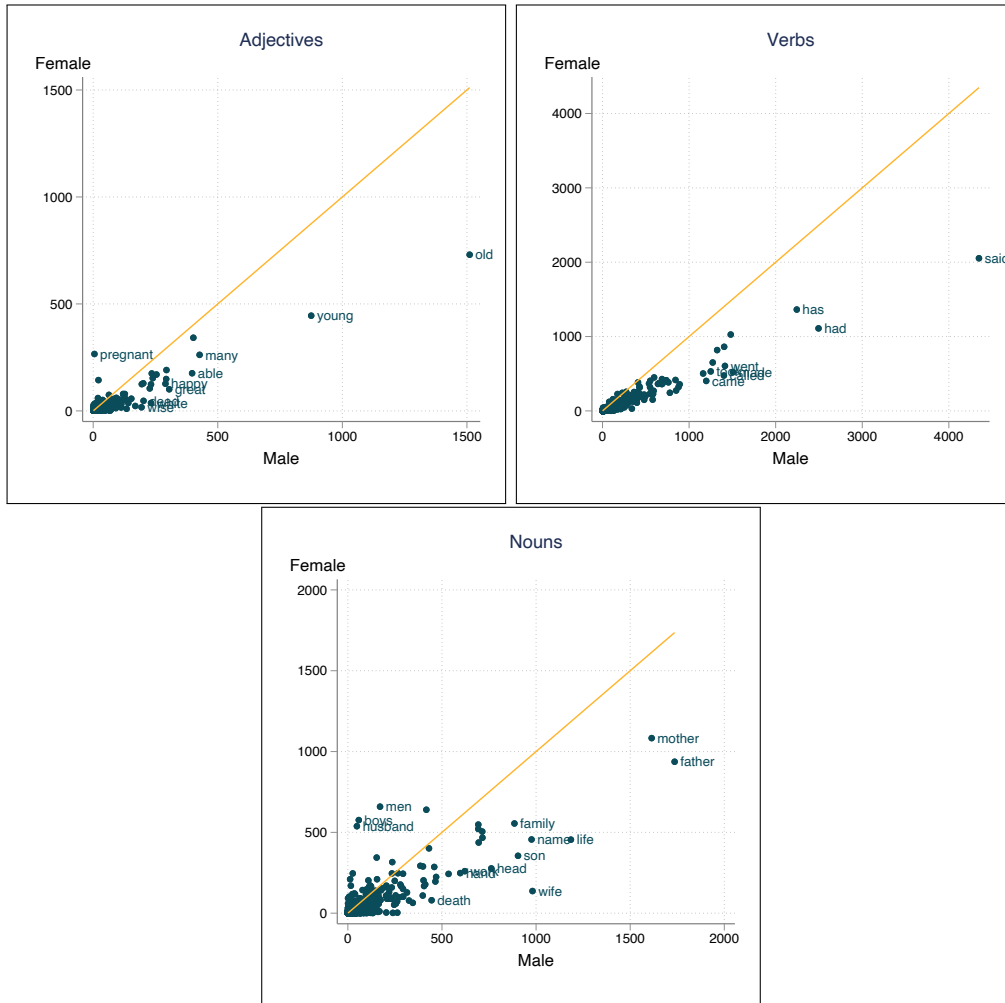
Note: This shows the results described above for the nine countries with over 250 occupation term-gender term cooccurrences.

Figure A6: Gender stereotypes, for countries with at least 1.5 million tokens



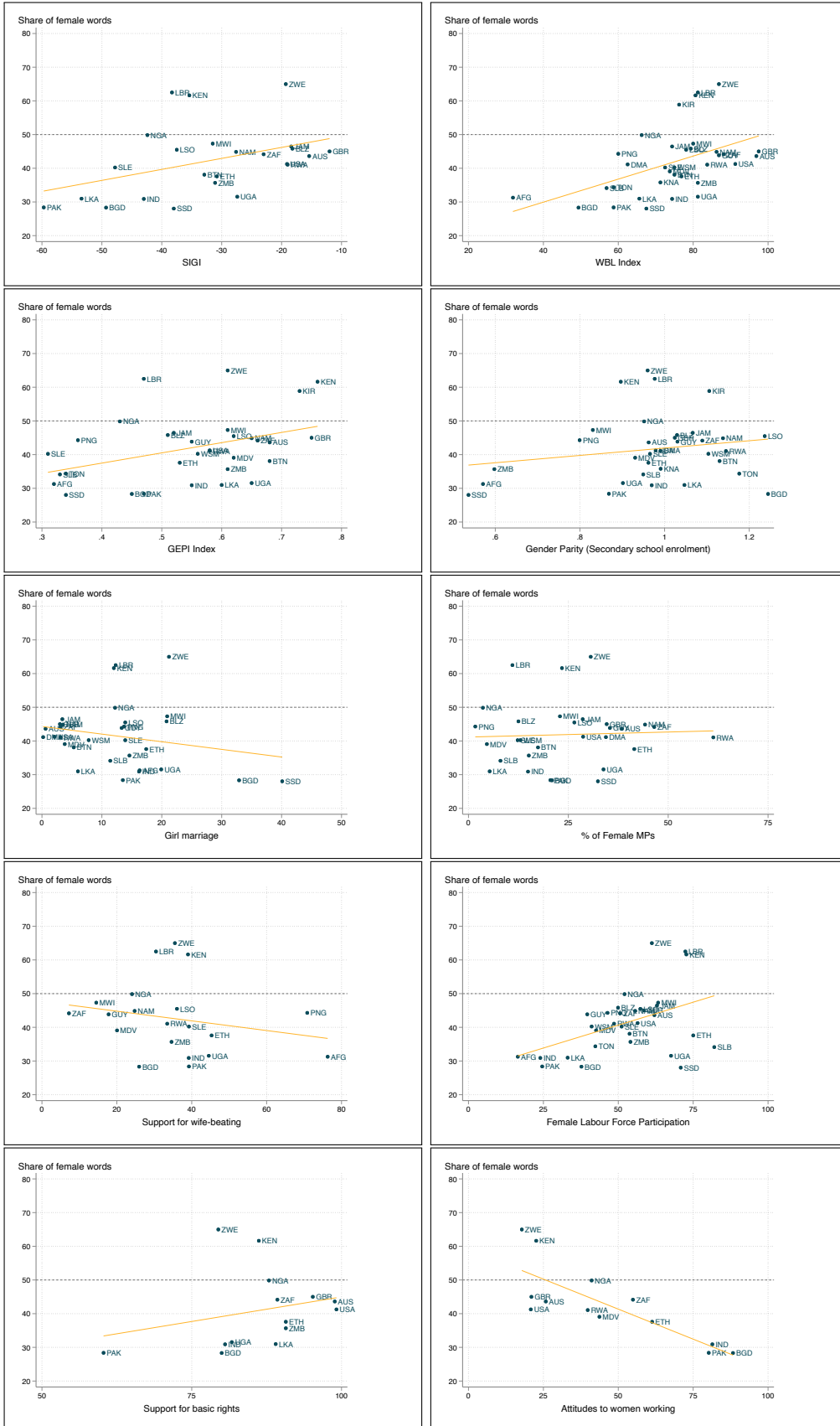
Note: Biases are shown only for the 10 countries in our corpus with at least 1.5 million tokens, as embeddings generated for other countries will be estimated unreliably. Male bias is calculated as the difference between the average cosine similarity of the theme word with the set of male gender terms, and the average similarity of the theme word with the set of female gender terms. Confidence intervals are calculated as the standard deviation for this statistic, over 50 bootstrap samples, where samples are generated by sampling all sentences in our corpus with replacement, until a sample has as many sentences as the original corpus.

Figure A7: Verbs, adjectives, and nouns used alongside gendered terms (Part-of-speech tagging)



Note: These figures show the counts of the number of occurrences of gendered words alongside different specific adjectives, verbs, and nouns. The gendered words used are those contained in Table C1 and the names described in Section 2.2. We identify adjectives, verbs, and nouns using the part-of-speech tagging approach described in Section 3.3.3.

Figure A8: Gender representation by country



Note: This figure shows correlations between gendered word representation and country characteristics. SIGI is the OECD Social Institutions & Gender Index. WBL is the WB Women, Business, & the Law Index. GEPI is CGD's Girl's Education Policy Index. Secondary Education Parity is from UNESCO. Girl marriage and female MPs are from the OECD Gender, Institutions, & Development database. Support for wife-beating is from the Demographic & Health Survey. Female labour force participation is from the WB World Development Indicators. Support for basic rights and attitudes to women working are from the Gallup World Poll.

B. Appendix: tables

Table B1: Publishing information for textbooks from each country

Country	Policy on Textbook Selection	Source Type	Publisher Type	External agency funding	Publisher
Afghanistan	Government-approved	Government website	Government		Ministry of Education
Australia	School Autonomy	Other	Private		Pascal Press
Bangladesh	Government-approved	Government website	Government		National Curriculum & Textbook Board
Belize	School Autonomy	Government website	Government		Ministry of Education
Bhutan	Government-approved	Government website	Government	Canadian International Development Agency	Ministry of Education
Dominica	School Autonomy	Government website	Government		Ministry of Education
Ethiopia	Government-approved	Government website	Government	Governments of Finland, Italy, Netherlands and the United Kingdom.	Ministry of Education
Guyana	School Autonomy	Government website	Government	Inter-American Development Bank	Ministry of Education
India	Mixed	Other	Government		National Council of Ed Research & Training
Jamaica	School Autonomy	Government website	Government		Ministry of Education
Kenya	Government-approved	Other	Government	USAID, RTI	Ministry of Education
Kiribati	Government-approved	Government website	Government		Ministry of Education
Lesotho	Government-approved	Other	Government	Commonwealth of Learning	Ministry of Education
Liberia	Government-approved	Other	Government	USAID	Ministry of Education
Maldives	Government-approved	Government website	Government		National Institute of Education, Cambridge University Press
Namibia	School Autonomy	Other	Government	Commonwealth of Learning	Ministry of Education
Nigeria	Government-approved	Other	Government	USAID	Ministry of Education
Pakistan	Mixed	Other	Government, Private		Ch. Ghulam Rasul and Sons, Punjab & Sindh Curriculum Boards
Papua New Guinea	Government-approved	Government website	Government	Japan	Department of Education
Rwanda	Government-approved	Government website	Government, Private		Rwanda Education Board, Longhorn
Samoa	Government-approved	Government website	Government	NZAID	Ministry of Education, Sports and Culture
Seychelles	Government-approved	Other	Government	Commonwealth of Learning	Ministry of Education
Sierra Leone	Government-approved	Government website	Government	UKAID; Cambridge Education; International Rescue Committee	Min of Basic & Senior Sec Ed
Solomon Islands	Government-approved	Government website	Government		Min of Ed & Human Resource Development
South Africa	School Autonomy	Both	Government	USAID	Department of Basic Education, Siyavula
South Sudan	Government-approved	Other	Government	GPE	Ministry of General Education & Instruction
Sri Lanka	Government-approved	Government website	Government		Ministry of Education
St Kitts and Nevis	School Autonomy	Government website	Government		Ministry of Education
Tonga	Government-approved	Government website	Government, Foreign agency		Secretariat of the Pacific Community, United States Department of Agriculture
Trinidad and Tobago	Government-approved	Other	Government	Commonwealth of Learning	Ministry of Education
Uganda	Government-approved	Government website	Government		National Curriculum Development Center
United Kingdom	School Autonomy	Other	Private		Collins
United States	School Autonomy	Other	Private		Macmillan McGraw
Zambia	Government-approved	Government website	Government	Commonwealth of Learning	Commonwealth of Learning, Curriculum Development Center
Zimbabwe	Government-approved	Government website	Government	Girl's Education Challenge - UKAID	Ministry of Primary & Secondary Education

Note: This table shows the sources of the textbooks included in our corpus.

C. Appendix: Word lists

Table C1: Gendered word lists

Female	Age	Male	Age	Female	Age	Male	Age
Family				Identities			
Mother	Old	Father	Old	Woman	Old	Man	Old
Mothers	Old	Fathers	Old	Women	Old	Men	Old
Aunt	Old	Uncle	Old	Female		Male	
Auntie	Old			Females		Males	
Aunties	Old			Girl	Young	Boy	Young
Aunts	Old	Uncles	Old	Girls	Young	Boys	Young
Aunty	Old			Characters			
Daughter	Young	Son	Young	Ladies		Gentlemen	Old
Daughters	Young	Sons	Young	Lady		Gentleman	Old
Grandma	Old	Granddad	Old			Knight	Old
Grandmas	Old	Granddads	Old	Madame	Old	Sirs	Old
Grandmom	Old	Grandpas	Old				
Grandmother	Old	Grandfather	Old	Maiden	Young	Lad	Young
Grandmothers	Old	Grandfathers	Old	Maidens	Young	Lads	Young
Granny	Old	Grandpa	Old			Master	Young
Mom	Old	Dad	Old			Masters	Young
Mommies	Old	Daddies	Old	Maids		Squires	Young
Mommy	Old	Daddy	Old	Princess	Young	Prince	Young
Moms	Old	Dads	Old	Princesses	Young	Princes	Young
Mum	Old	Papa	Old	Queen	Old	King	Old
Mummies	Old	Papas	Old	Queens	Old	Kings	Old
Mummy	Old			Witch		Wizard	
Mums	Old			Witches		Wizards	
Stepmother	Old	Stepfather	Old			Chap	
Stepmothers	Old	Stepfathers	Old			Chaps	
Wife	Old	Husband	Old	Actress		Actor	
Wives	Old	Husbands	Old	Waitress		Waiter	
Nana	Old			Maid		Squire	Young
Nanas	Old					Fellow	Old
Sister		Brother				Fellows	Old
Sisters		Brothers		Damsel	Young	Paige	Young
		Bro		Damsels	Young	Paiges	Young
Pronouns				Emperess	Old	Emperor	Old
Her		Him		Emperesses	Old	Emperors	Old
Hers		His		Empress	Old	Lord	Old
Herself		Himself		Empresses	Old	Lords	Old
		Hisself					
Miss	Young						
Mrs	Old	Mr	Old				
Ms							
She		He					

Table C2: Occupation word lists

Professional and managerial		Service workers	Manual workers
accountant	lawyer	attendant	baker
actor	lecturer	barber	blacksmith
administrator	librarian	butler	builder
architect	magician	carpenter	driver
artist	magistrate	cashier	farmer
astronomer	mathematician	cleaner	gardener
biologist	merchant	clerk	goldsmith
buyer	musician	dealer	hawker
chemist	nurse	housekeeper	labourer
consultant	photographer	operative	mechanic
counsellor	physician	operator	messenger
doctor	physicist	receptionist	miller
economist	pilot	secretary	porter
editor	preacher	servant	potter
engineer	priest	steward	shepherd
envoy	producer	tailor	soldier
executive	proprietor	tutor	
governor	publisher	waiter	
historian	reporter	weaver	
illustrator	researcher		
inspector	scientist		
instructor	sculptor		
interpreter	supervisor		
interviewer	surgeon		
inventor	surveyor		
journalist	translator		
judge	treasurer		
landlord	writer		

Table C3: Theme word lists

Achievement	Appearance	Family	Work
Achievement	Alluring	Children	Business
Authority	Elegant	Chore	Career
Better	Beautiful	Cousins	Company
Control	Slim	Domestic	Corporation
Effort	Ugly	Family	Economy
Leader	Gorgeous	Home	Executive
Plan		Household	Industry
Persuasive		Marriage	Job
Power		Parents	Labor
Powerful		Relatives	Labour
Success		Wedding	Management
Tried			Office
Won			Pay
			Professional
			Salary
			Trade
			Wage
			Work
			Workers
			Working

D. Appendix: Example textbook pages

Figure D1: Sample textbook page from Sierra Leone, Grade 10 English

Practice

Activity 1

Identify which of the spelling rules apply to the words in **bold**.

1. The **ladies** meet at the tea shop every Thursday.
2. Put the books on the **shelves**.
3. They went **shopping** in Freetown.
4. **He** has 4 **donkeys** on **his** smallholding.
5. **He** is **writing** an essay for **his** homework.
6. **He** **stopped** playing football because **he** had to study.
7. Mosquitoes were **biting** me last night.
8. **She** saw **her** friend last night.
9. The **leaves** are falling off the tree because it is getting cold.
10. Did you **receive** a letter from your grandparents?

Activity 2

Correct the spelling in the following paragraph.

Freetown is a great place to go shopping. The shelves in the supermarkets are full. You would not believe how much there is in the markets. I often go out with friends who are living there. They are my grandparents' neighbours. After shopping, we like sitting and looking at all our purchases. I think it is the same all over the world – **ladys** like to shop.

Note: In this sample page from a textbook we can see an uneven representation of male and female nouns and pronouns - there are 6 male themed words (he/his) and 4 female themed words (ladies/she/her/ladys).

Figure D2: Sample textbook page from Pakistan, Grade 5 Social Studies

Federal System of Government **CHAPTER 5**

National Assembly (MNAs). They are elected for five years. The Speaker presides over the sessions of National Assembly. Important national and international affairs are discussed in the National Assembly.

Prime Minister

Prime Minister looks after the administrative affairs of the country. There are several Federal departments that run the system of the country. Each department is headed by a Federal Secretary who works under the supervision of Federal Minister.

According to the Constitution, National Assembly is elected after every five years. Members of National Assembly are elected in these elections. Prime Minister is elected by the members of the party which is in majority. If a party does not have a clear majority in the assembly, then different parties make alliance with one another and elect the Prime Minister. Prime Minister has his own Secretariat.

Cabinet

Prime Minister has to fulfil many national and international responsibilities. He cannot handle the affairs of the country all alone. He appoints ministers for his help. The number of ministers is decided by him according to the Constitution. All the Federal Ministers constitute Federal Cabinet. They are made heads of different departments like commerce, finance etc. To run administration of a department, a state official is appointed as secretary. There are many government officials under the Secretary. The meetings of Federal Cabinet are called in which important policy decisions are made through mutual consultation. Performance of different government departments is also evaluated in such meetings.

42



Session of National Assembly



Prime Minister House, Islamabad

Note: In this sample page we see the Prime Minister being discussed as He/him/his. The current Prime Minister does happen to be a man, but he is not named in this book, and the role is clearly not inherently gendered.

Figure D3: Sample textbook page from Pakistan, Grade 7 Home Economics



Note: This book discusses how “you can help your mother” with housework.

Example sentences of gender stereotypes

Afghanistan

Grade: 12

Subject: English

1. Layla started the washing-machine an hour ago. There are now ten clean pants in the cupboard (p28)
2. Captain Abdullah is the Ariana's airline pilot. He started flying for Afghanistan twenty years ago. (p28)
3. She didn't make that cake herself (p46)

Grade: 10

Subject: English

1. She prays five times a day (p52)
2. He works hard (p52)
3. She had better not stay out late (p56)
4. She didn't sew the curtain herself (p82)

Grade: 7

Subject: English

1. My name is Farhad my father name is Qadeem. My father is a doctor. (p33)
2. His father is an engineer (p36)
3. Our aunt has a garden (p42)
4. Their brother has a building (p42)
5. My father has a factory (p42)
6. Her mother has a brother (p42)

Bangladesh

Grade: 7

Subject: English

1. Everyday Shahana helps with the household work (p28)

Pakistan

Grade: 5

Subject: English

1. "I have prepared our small supper," said his wife (p11)
2. He was an intelligent boy and his mother wished him to study (p17)
3. My mother is a great cook (p25)
4. He was an efficient and hardworking student (p31)
5. She made dinner for the whole family (p41)
6. He fell while catching a ball and hurt his knees (p72)

Grade: 8

Subject: Math

1. Find the principal amount invested by Riaz in a business if he receives a profit of Rs. 4200 in 3 years at the rate of 10% per year. (p25)* All word problems involving wealth/purchasing items/insurance in the textbook are male subjects
2. Saeeda has 196 marbles that she is using to make a square formation. How many marbles should be in each row? (p27)* only word problem with a female subject

Grade: 9

Subject: English

1. She prepared food for this journey (p33)
2. When Aslam was playing cricket, he hurt his hand (p39)
3. My mother always prays for my success (p131)

4. My great-grandfather owned a big house in his hometown (p132)

Sierra Leone

Grade: 10

Subject: English

1. I (she) helped my children look after their children, and then my grandchildren after that (p30)
2. He had a very important job interview at an office in town (p35)
3. He studied every day when he got home from school for a couple of hours (p48)
4. Samuel got a scholarship to college, so his family did not have to worry about finding the money (p49)
5. On Wednesday afternoon, she is going to stay home (p67)
6. He worked so hard at school that he won a scholarship to the university to study medicine (p75)
7. The ladies meet at the tea shop every thursday (p83)
8. I think it is the same all over the world - ladys like to shop (p83 misspelling intentional)

Sri Lanka

Grade: 4

Subject: English

1. Senuri keeps her garden clean. She sweeps the garden. She picks leaves. She waters the plants. She dumps garbage into the bin. She weeds the garden (p13)
2. I'm making sweetmeats with my mother (p46)
3. I'm going to the mosque with my father (p46)
4. Senuri is drinking tea with her sister (p77)