What might books be teaching young children about gender?

XXX[1], XX[2], XXX[2], XXX[2], & XXX[2]

[1] XXX

[2] XXX

Abstract

We investigated how gender is represented in children's books using a 200,000 word corpus comprising 247 popular, contemporary books for young children (0-5 years). Using human judgments and word co-occurrence data, we quantified the gender biases of words within the corpus and within individual books. We find that children's books contain large numbers of words that adults judge as gendered. Semantic analyses based on co-occurrence data yielded word clusters related to gender stereotypes (e.g., feminine: emotions; masculine: tools). Co-occurrence data also indicate that books instantiate gender stereotypes found in other research (e.g., girls are better at reading and boys at math). Finally, we used large-scale data to estimate the gender distribution of the audience for individual books, and find that children tend to be exposed to gender stereotypes for their own gender. Together the data suggest that children's books may be an early source of gender stereotypes.

STATMENT OF RELEVANCE: Gender stereotypes such as that girls are better at reading and boys are better at math have origins in early childhood. We examined the extent to which popular children's books may expose young 0-5 year old children to information about gender stereotypes. Our data suggest that children's books have rich information about gender stereotypes, and that some of these stereotypes are more strongly represented in children's books than in adult fiction. These findings suggest that popular children's books may be unintentionally teaching young children about gender stereotypes.

Keywords: reading, gender, language development

Word count: 2051 (excluding methods and results)

What might books be teaching young children about gender?

Table 1
Examples of Clusters from Multi-Dimensional Embeddings

Category	Effect Size	*N*	Examples
Female-Biased Clusters			
affection	1.33 [0.9, 2.1]	21	kisses, loved, smile, tears, heart, care
modifiers	0.79 [0.49, 1.27]	34	probably, whenever, truly, likely, completely, yet
communication verbs	0.74 [0.43, 1.14]	25	spoke, listened, heard, explained, asked, answered
school	0.54 [0.12, 1.12]	20	learning, practicing, school, students, writing, book
food	0.44 [0.15, 0.8]	43	meatballs, soup, eggs, milk, pie, salad
Neutral Clusters			
family relationships	0.19 [-0.18, 0.63]	29	children, brother, sister, uncle, aunt
body parts	0.14 [-0.16, 0.48]	41	eye, knee, ankle, hair, bone
house parts	0.08 [-0.24, 0.4]	40	bedroom, floor, lamp, roof, window
quantifiers	0.05 [-0.29, 0.4]	36	few, almost, many, most, whole
spatial terms	-0.31 [-0.71, 0.02]	39	across, long, low, through, close
Male-Biased Clusters			
zoo animals	-0.53 [-1.27, -0.07]	23	giraffe, elephant, gorilla, lion, monkey, zebra
airborne actions	-0.83 [-1.21, -0.54]	37	climbed, tossed, jumped, knocked, pulled, swung
tools	-0.89 [-1.42, -0.52]	20	axe, blade, knife, bow, stick, wood
transportation (ground)	-1.23 [-1.62, -0.93]	40	car, bicycle, trains, ambulance, engine, traffic
professions	-1.35 [-2.19, -0.92]	23	judge, policemen, guard, sailor, mayor, clerk

Note: Effect size measure is Cohen's *d* based on a one-sample *t*-test comparing the mean gender of words in a cluster to the overall word gender mean. Clustering is an unsupervised machine learning method for dividing observations into k clusters by minimizing within-cluster distance and maximizing across-cluster distance. Brackets give bootstrapped 95 percent confidence intervals. *N* indicates number of words in each cluster.

Table 2
Four IATs used to study gender bias

Psychological Bias	Target Words	Behavioral Studie
women as good; men as bad	"good": good, happy, gift, sunshine, heaven "bad": bad, awful, sick, trouble, hurt	Cveneck, Meltzoff, \& G
		Skowronski \& Lawrenc
		Greenwald et al. (2002,
		Goodman (2004, A)
women and family; men and career	"family": family, parents, children, home, cousins, wedding "career": job, work, money, office, business, desk	Nosek, Banaji, \& Gree
women and language; men and math	"language": books, read, write, story, letters, spell "math": numbers, count, sort, size, shapes, different	Cveneck, Meltzoff, Gree
		Nosek, Banaji, \& Gree

women and arts; men and math	"art": art, paint, draw, books, dance, story "math": numbers, count, sort, size, shapes, different	Nosek, Banaji \& Green

car	mpg	cyl
Mazda RX4	21	6
Mazda RX4 Wag	21	6
Datsun 710	22.8	4
Hornet 4 Drive	21.4	6
Hornet Sportabout	18.7	8
Valiant	18.1	6
Duster 360	14.3	8
Merc 240D	24.4	4
Merc 230	22.8	4
Merc 280	19.2	6