SUPPLEMENTARY ONLINE MATERIALS

Evidence of Bias against Girls and Women

in Contexts that Emphasize Intellectual Ability

Table of Contents

Table S1

The Job Descriptions Used in Experiments 1 and 2

Experimental condition
We are looking for applicants to fill two new positions in our workforce. These positions will involve a lot of responsibility and the opportunity to join one of our newest, and most exciting, departments. Because of the work we do, we are looking for candidates who have a high IQ, superior reasoning skills, and a knack for big, bold ideas. That is, we'd like to hire someone whose intellectual abilities stand out from those of their peers. Our work environment values and emphasizes employees' natural intelligence, expecting everyone to push their inborn smarts to the limit. Therefore, we are especially interested in candidates who demonstrate an inherent aptitude for this position.
Control condition
We are looking for applicants to fill two new positions in our workforce. These positions will involve a lot of responsibility and the opportunity to join one of our newest, and most exciting, departments. Because of the work we do, we are looking for highly motivated candidates with an outstanding work ethic and a superior commitment to doing their work as well as possible. That is, we'd like to hire someone who has demonstrated significant and sustained dedication in their past positions. Our work environment values and emphasizes employees' strivings and their consistent effort to achieve goals, expecting everyone to continuously improve their work performance. Therefore, we are especially interested in candidates who demonstrate continual passion for the job.

Appendix S1

Additional Measures in Experiments 1 and 2

Aside from the measures described in the main text, Experiments 1 and 2 included several additional measures. (i) Participants indicated *why* they referred the person they did. A cursory inspection of these justifications indicated that the vast majority of participants either stated that the person fit the job description or listed the traits that had been emphasized in the job description (e.g., "he is incredibly smart," "she told me she is a member of Mensa"). (ii) Participants indicated how much each referred individual should be paid if offered the job. Exploratory analyses revealed no differences between the salaries offered to the men and women referred in the two conditions. (iii) Participants filled out the measure of stereotyping a second time, from *society's* perspective ("Please rate to rate the extent to which society at large agrees with these statements"). Perhaps not surprisingly, this measure was not predictive of participants' referrals, unlike the measure of their personal endorsement of stereotypes about the intellectual abilities of men and women. (iv) After the measure of stereotyping, we administered the Modern Sexism Scale (e.g., "Discrimination against women is no longer a problem in the United States"; 1 = *strongly disagree* to 9 = *strongly agree*; Swim, Aikin, Hall, & Hunter, 1995). This measure was positively correlated with our measure of stereotyping and—like it—was also inversely related to the probability of referring a woman. (v) At the end of the survey, participants were asked to indicate how prestigious they believed the job to be (1 = *not at all prestigious* to 9 = *extremely prestigious*). Participants rated the "brilliance" job to be more prestigious, but their prestige ratings did not predict the gender of their referrals, nor did statistically adjusting for prestige diminish the effect of our manipulation.

Table S2

The Bayesian Mixed-effects Logistic Regression Model in Experiment 1

	b	*SE*	95% Credible Interval	
Intercept	-0.1355	0.0830	-0.2977	0.0267
Condition (0 = control; 1 = experimental)	-0.4834	0.1640	-0.8066	-0.1651
Gender (0 = man; 1 = woman)	0.9665	0.1670	0.6350	1.3009
Stereotype	-0.0873	0.0473	-0.1806	0.0071
Condition x Gender	-0.1468	0.3271	-0.7903	0.4899
Condition x Stereotype	-0.0800	0.0918	-0.2606	0.1000
Gender x Stereotype	0.0053	0.0955	-0.1832	0.1929
Condition x Gender x Stereotype	-0.2740	0.1863	-0.6369	0.0844

Note. The dependent variable was the gender of the referred person (0 = man; 1 = woman). All variables were mean-centered prior to inclusion in the model.

Table S3

The Bayesian Mixed-effects Logistic Regression Model in Experiment 2

	b	*SE*	95% Credible Interval	
Intercept	-0.0956	0.0556	-0.2021	0.0127
Condition (0 = control; 1 = experimental)	-0.2224	0.1097	-0.4438	-0.0067
Gender (0 = man; 1 = woman)	1.1965	0.1127	0.9753	1.4173
Ethnicity (0 = participant of color; 1 = white participant)	-0.3334	0.1119	-0.5500	-0.1163
Stereotype	-0.0510	0.0298	-0.1091	0.0082
Condition x Gender	-0.3647	0.2186	-0.7861	0.0650
Condition x Ethnicity	0.2431	0.2176	-0.1796	0.6742
Gender x Ethnicity	-0.0382	0.2297	-0.4850	0.4032
Condition x Stereotype	0.0411	0.0579	-0.0735	0.1538
Gender x Stereotype	0.1170	0.0592	0.0010	0.2334
Ethnicity x Stereotype	0.0109	0.0584	-0.1052	0.1248
Condition x Gender x Ethnicity	-0.4784	0.4207	-1.2946	0.3534
Condition x Gender x Stereotype	-0.1558	0.1144	-0.3816	0.0679
Condition x Ethnicity x Stereotype	-0.0663	0.1138	-0.2925	0.1617
Gender x Ethnicity x Stereotype	0.1558	0.1167	-0.0736	0.3864
Condition x Gender x Ethnicity x Stereotype	-0.2798	0.2286	-0.7245	0.1711

Note. The dependent variable was the gender of the referred person (0 = man, 1 = woman). All variables were mean-centered prior to inclusion in the model.

Table S4

The Unfamiliar Games Used in Experiment 3

Zorb I want to tell you about this game that I ask children to play sometimes. It's called Zorb, and it's a lot of fun. In this game, what you have to do is to figure out how to get two small coins and two big coins in each green tube. But you also need to make sure to have all the big ones at the bottom and all the small ones at the top without using any of the pink tubes more than once. [*Oh, and here is something else about the Zorb game, and this is important, so make sure you're paying attention. This game is not for everyone. It's only for children who are really, really smart. Only smart children can be good at this game.*]	
Tever I want to tell you about this game that I ask children to play sometimes. It's called Tever, and it's a lot of fun. In this game, what you have to do is to figure out how to put as many pencils and erasers as you can into these holes in just one minute. But you also need to have the pencils in a straight line and the erasers in a triangle shape without having the straight lines cross the triangles. [*Oh, and here is something else about the Tever game, and this is important, so make sure you're paying attention. This game is not for everyone. It's only for children who are really, really smart. Only smart children can be good at this game.*]	

Note. All children heard the first part of each game description. Only children in the experimental condition heard the italicized text in square brackets.

Table S5

The Bayesian Mixed-effects Logistic Regression Model in Experiment 3

	b	*SE*	95% Credible Interval	
Intercept	0.1671	0.1250	-0.0798	0.4150
Condition (0 = control; 1 = experimental)	-0.2253	0.2419	-0.6994	0.2463
Gender (0 = boy; 1 = girl)	2.3654	0.2545	1.8789	2.8714
Age (5 to 7 years; continuous)	-0.3053	0.1424	-0.5878	-0.0259
Stereotype (0 to 1; continuous)	-1.4301	0.4576	-2.3232	-0.5423
Selection Round (1, 2, or 3)	-0.3531	0.0985	-0.5485	-0.1615
Condition x Gender	-0.7036	0.4417	-1.5606	0.1598
Condition x Age	0.0824	0.2714	-0.4473	0.6175
Gender x Age	-0.0255	0.2813	-0.5636	0.5294
Condition x Stereotype	0.4668	0.7171	-0.9463	1.8696
Gender x Stereotype	-0.1852	0.7189	-1.5952	1.2150
Age x Stereotype	0.8575	0.5103	-0.1603	1.8533
Condition x Selection Round	-0.4167	0.1946	-0.8054	-0.0372
Gender x Selection Round	-0.6165	0.1899	-0.9905	-0.2443
Age x Selection Round	0.2175	0.1144	-0.0042	0.4451
Stereotype x Selection Round	0.5716	0.3881	-0.1697	1.3378
Condition x Gender x Age	-0.4855	0.4994	-1.4646	0.5007
Condition x Gender x Stereotype	0.2746	0.8918	-1.5175	2.0123
Condition x Age x Stereotype	-0.6408	0.7609	-2.1126	0.8706
Gender x Age x Stereotype	0.2796	0.7545	-1.2069	1.7501
Condition x Gender x Selection Round	-0.1985	0.3683	-0.9161	0.5277
Condition x Age x Selection Round	-0.3118	0.2187	-0.7375	0.1135
Gender x Age x Selection Round	-0.2608	0.2209	-0.6921	0.1764
Condition x Stereotype x Selection Round	-0.1460	0.6346	-1.4054	1.1057
Gender x Stereotype x Selection Round	-0.7644	0.6430	-2.0274	0.4933
Age x Stereotype x Selection Round	0.8134	0.4238	-0.0082	1.6407
Condition x Gender x Age x Stereotype	0.6151	0.9179	-1.1918	2.4096
Condition x Gender x Age x Selection Round	0.3543	0.4147	-0.4595	1.1692
Condition x Gender x Stereotype x Selection Round	0.0713	0.8687	-1.6610	1.7464
Condition x Age x Stereotype x Selection Round	0.9142	0.6894	-0.4238	2.2638
Gender x Age x Stereotype x Selection Round	-0.4843	0.6895	-1.8464	0.8609
Condition x Gender x Age x Stereotype x Selection Round	0.1752	0.9004	-1.5904	1.9821

Note. The dependent variable was the gender of the selected teammate (0 = boy, 1 = girl). All variables were mean-centered prior to inclusion in the model.

Table S6

The Maximum-Likelihood Mixed-effects Logistic Regression Model in Experiment 1

	b	*SE*	*z*	*p*
Intercept	-0.1313	0.0807	-1.626	0.1038
Condition (0 = control; 1 = experimental)	-0.4879	0.1616	-3.018	0.0025
Gender (0 = man; 1 = woman)	0.9700	0.1664	5.828	5.63e-09
Stereotype	-0.0831	0.0460	-1.806	0.0709
Condition x Gender	-0.1580	0.3333	-0.474	0.6352
Condition x Stereotype	-0.0803	0.0921	-0.872	0.3833
Gender x Stereotype	0.0036	0.0929	0.040	0.9684
Condition x Gender x Stereotype	-0.2806	0.1859	-1.509	0.1312

Note. The dependent variable was the gender of the referred person (0 = man; 1 = woman). All variables were mean-centered prior to inclusion in the model. The R syntax for the model was as follows: `glmer(nominated_woman ~ brilliance_c*female_c*stereotype_c + (1|subj), data=[data], family=binomial).`

Table S7

The Maximum-Likelihood Mixed-effects Logistic Regression Model in Experiment 2

	b	*SE*	*z*	*p*
Intercept	-0.0929	0.0555	-1.674	0.0941
Condition (0 = control; 1 = experimental)	-0.2243	0.1109	-2.022	0.0431
Gender (0 = man; 1 = woman)	1.1905	0.1125	10.583	<2e-16
Ethnicity (0 = participant of color; 1 = white participant)	-0.3322	0.1116	-2.977	0.0029
Stereotype	-0.0497	0.0288	-1.726	0.0844
Condition x Gender	-0.3796	0.2248	-1.688	0.0913
Condition x Ethnicity	0.2513	0.2229	1.128	0.2595
Gender x Ethnicity	-0.0308	0.2269	-0.136	0.8917
Condition x Stereotype	0.0394	0.0574	0.687	0.4923
Gender x Stereotype	0.1160	0.0577	2.010	0.0443
Ethnicity x Stereotype	0.0110	0.0576	0.191	0.8486
Condition x Gender x Ethnicity	-0.5680	0.4534	-1.253	0.2103
Condition x Gender x Stereotype	-0.1580	0.1150	-1.374	0.1695
Condition x Ethnicity x Stereotype	-0.0706	0.1149	-0.615	0.5386
Gender x Ethnicity x Stereotype	0.1595	0.1156	1.379	0.1678
Condition x Gender x Ethnicity x Stereotype	-0.2927	0.2307	-1.269	0.2044

Note. The dependent variable was the gender of the referred person (0 = man, 1 = woman). All variables were mean-centered prior to inclusion in the model. The R syntax for the model was as follows: `glmer(nominated_woman ~ brilliance_c*female_c*white_c*stereotype_c + (1|subj), data=[data], family=binomial).`

Table S8

The Maximum-Likelihood Mixed-effects Logistic Regression Model in Experiment 3

	b	*SE*	*z*	*p*
Intercept	0.1699	0.1187	1.431	0.1524
Condition (0 = control; 1 = experimental)	-0.1800	0.2377	-0.757	0.4488
Gender (0 = boy; 1 = girl)	2.3506	0.2525	9.309	<2e-16
Age (5 to 7 years; continuous)	-0.3196	0.1368	-2.336	0.0194
Stereotype (0 to 1; continuous)	-1.6893	0.5140	-3.287	0.0010
Selection Round (1, 2, or 3)	-0.3654	0.0995	-3.671	0.0002
Condition x Gender	-0.8721	0.4757	-1.833	0.0667
Condition x Age	0.1740	0.2731	0.637	0.5238
Gender x Age	0.0130	0.2730	0.048	0.9618
Condition x Stereotype	1.0426	1.0175	1.025	0.3055
Gender x Stereotype	-0.2729	1.0145	-0.269	0.7878
Age x Stereotype	1.0625	0.5687	1.868	0.0617
Condition x Selection Round	-0.4134	0.1983	-2.084	0.0371
Gender x Selection Round	-0.6121	0.1986	-3.081	0.0020
Age x Selection Round	0.2089	0.1145	1.823	0.0682
Stereotype x Selection Round	0.6464	0.4367	1.480	0.1388
Condition x Gender x Age	-0.6592	0.5465	-1.206	0.2277
Condition x Gender x Stereotype	1.5463	2.0295	0.762	0.4461
Condition x Age x Stereotype	-1.4236	1.1356	-1.254	0.2099
Gender x Age x Stereotype	0.3912	1.1336	0.345	0.7300
Condition x Gender x Selection Round	-0.1986	0.3964	-0.501	0.6162
Condition x Age x Selection Round	-0.2884	0.2289	-1.260	0.2077
Gender x Age x Selection Round	-0.2535	0.2290	-1.107	0.2681
Condition x Stereotype x Selection Round	-0.1515	0.8722	-0.174	0.8620
Gender x Stereotype x Selection Round	-1.2210	0.8728	-1.399	0.1618
Age x Stereotype x Selection Round	0.9281	0.4918	1.887	0.0591
Condition x Gender x Age x Stereotype	3.8237	2.2714	1.683	0.0923
Condition x Gender x Age x Selection Round	0.4474	0.4579	0.977	0.3285
Condition x Gender x Stereotype x Selection Round	0.2816	1.7446	0.161	0.8717
Condition x Age x Stereotype x Selection Round	1.7210	0.9820	1.752	0.0796
Gender x Age x Stereotype x Selection Round	-0.9514	0.9837	-0.967	0.3334
Condition x Gender x Age x Stereotype x Selection Round	0.8205	1.9647	0.418	0.6762

Note. The dependent variable was the gender of the selected teammate (0 = boy, 1 = girl). All variables were mean-centered prior to inclusion in the model. The R syntax for the model was as follows: `glmer(chose_girl ~ brilliance_c*female_c*age_c*stereotype_c*trial_c + (1|subj), data=[data], family=binomial, control = glmerControl(optimizer = "bobyqa", optCtrl=list(maxfun=1e5)), nAGQ = 10).` The change to the default optimizer settings in the *glmer* syntax was made so that the model can converge.

Table S9

The Bayesian Mixed-effects Logistic Regression Model in Experiment 3 (Boys' Data Only)

	b	*SE*	95% Credible Interval	
Intercept	-1.2379	0.2336	-1.7189	-0.8080
Condition (0 = control; 1 = experimental)	0.2181	0.3930	-0.5597	0.9994
Age (5 to 7 years; continuous)	-0.2867	0.2391	-0.7657	0.1752
Stereotype (0 to 1; continuous)	-0.9423	0.6808	-2.2667	0.4156
Selection Round (1, 2, or 3)	-0.0355	0.1409	-0.3071	0.2440
Condition x Age	0.3308	0.4451	-0.5487	1.1955
Condition x Stereotype	0.0501	0.8752	-1.6563	1.7548
Age x Stereotype	0.4438	0.7133	-0.9770	1.8373
Condition x Selection Round	-0.3061	0.2719	-0.8328	0.2352
Age x Selection Round	0.3860	0.1656	0.0635	0.7052
Stereotype x Selection Round	1.0402	0.5271	0.0190	2.0881
Condition x Age x Stereotype	-0.7102	0.9092	-2.4852	1.0876
Condition x Age x Selection Round	-0.5478	0.3151	-1.1750	0.0647
Condition x Stereotype x Selection Round	-0.1977	0.7863	-1.7348	1.3575
Age x Stereotype x Selection Round	1.0813	0.5911	-0.0685	2.2338
Condition x Age x Stereotype x Selection Round	0.4591	0.8131	-1.1556	2.0483

Note. The dependent variable was the gender of the selected teammate (0 = boy, 1 = girl). All variables were mean-centered prior to inclusion in the model.

Table S10

The Bayesian Mixed-effects Logistic Regression Model in Experiment 3 (Girls' Data Only)

	b	*SE*	95% Credible Interval	
Intercept	1.3059	0.1635	1.0094	1.6443
Condition (0 = control; 1 = experimental)	-0.5705	0.2800	-1.1357	-0.0415
Age (5 to 7 years; continuous)	-0.2940	0.1684	-0.6297	0.0349
Stereotype (0 to 1; continuous)	-1.2525	0.5214	-2.2893	-0.2371
Selection Round (1, 2, or 3)	-0.6292	0.1421	-0.9120	-0.3593
Condition x Age	-0.1702	0.3182	-0.7979	0.4455
Condition x Stereotype	0.5981	0.7811	-0.9511	2.1014
Age x Stereotype	0.9182	0.5581	-0.1850	2.0139
Condition x Selection Round	-0.4767	0.2750	-1.0239	0.0618
Age x Selection Round	0.0697	0.1579	-0.2365	0.3783
Stereotype x Selection Round	-0.0494	0.5127	-1.0402	0.9635
Condition x Age x Stereotype	0.1569	0.8069	-1.4554	1.7215
Condition x Age x Selection Round	-0.0883	0.3149	-0.7063	0.5282
Condition x Stereotype x Selection Round	-0.0155	0.7540	-1.4996	1.4680
Age x Stereotype x Selection Round	0.3439	0.5492	-0.7258	1.4127
Condition x Age x Stereotype x Selection Round	0.7551	0.7814	-0.7797	2.2753

Note. The dependent variable was the gender of the selected teammate (0 = boy, 1 = girl). All variables were mean-centered prior to inclusion in the model.