Appendix

A. Acknowledgments

I am grateful to my advisers, Steven Haider, Michael Conlin, Scott Imberman, and Kris Renn for their invaluable feedback. I am also grateful to Dylan Brewer, Michigan State University's Department of Economics graduate student workshop participants, and the Missouri Valley Economic Association's 2008 annual conference participants for their comments and suggestions.

B. Creation of analytic sample and variable definitions

The Add Health in-school questionnaire contains responses from 90,118 students, but I am not able to use all of these responses in my analysis. One student is missing a valid school identifier, 803 students are missing a valid grade, 6,849 students are missing a survey weight, and 17,325 students did not provide a valid answer to at least one of the outcomes or control variable questions. I also exclude 3,767 students who attended schools that required explicit parental consent to take the in-school questionnaire. These schools tend to have lower in-school questionnaire participation rates which makes the calculation of the share of female peers more prone to measurement error. Finally, I exclude 100 students who had fewer than ten total students in their grade and 1,144 students who attended schools that were exclusively male or female, as there is no variation in peer gender across grades. I am left with 60,129 observations.

My explanatory variable of interest is the share of female peers in a student's school and grade which I calculate by dividing the number of students who answered the question, "what sex are you?" with "female" by the total number of students who answered the question either "male" or "female" in that student's school and grade. I calculate the share of female peers from the full sample of in-school questionnaire respondents before I remove students for my analytic sample. To avoid correlations between the student's own gender and the share of female peers in their school and grade, I use a leave-out mean, where I do not include the student in the calculation of their peer group's mean.

I use the following survey questions as individual controls:

- Biological sex
- Hispanic of Spanish ethnicity
- Black or African American and not Hispanic
- Asian or Pacific Islander and not Hispanic
- American Indian or Native American and not Hispanic
- White and not Hispanic
- Other race not listed and not Hispanic
- If the student's age (in years) is above the median age for their grade across all schools
- If the student reported being born in the United States
- Dummy variables for the number of people that live in the student's household: 1, 2, 3, 4-6, and if the student lives in a shelter or group home

C. Descriptive statistics

The below table provides the unweighted and weighted mean of the share of female peers and control variables in the baseline model.

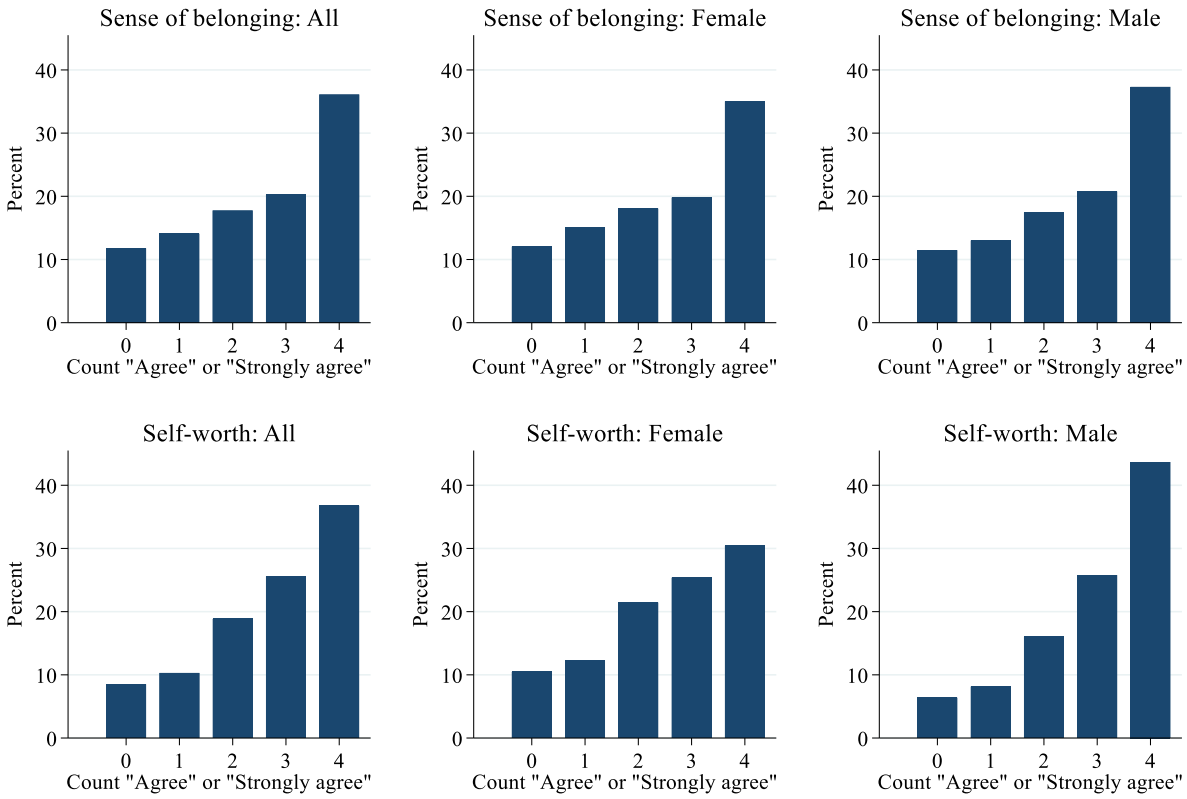
Table A1: Descriptive statistics

	Unweighted		Weighted	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Share female	0.503	0.045	0.503	0.054
Female	0.523	0.499	0.516	0.500
Hispanic or Latino	0.142	0.349	0.107	0.309
Black or African American	0.159	0.366	0.170	0.376
Asian or Pacific Islander	0.059	0.236	0.041	0.199
American Indian or Native American	0.033	0.178	0.037	0.189
Other race	0.024	0.154	0.023	0.150
Above median age	0.338	0.473	0.339	0.474
Born outside the United States	0.086	0.280	0.059	0.235
Household size				
One	0.005	0.071	0.005	0.073
Two	0.053	0.223	0.053	0.224
Three	0.183	0.387	0.185	0.388
Four to six	0.756	0.429	0.754	0.431
Shelter or group home	0.003	0.057	0.003	0.056
Observations	60,129		60,129	

D. Histograms of outcome variables

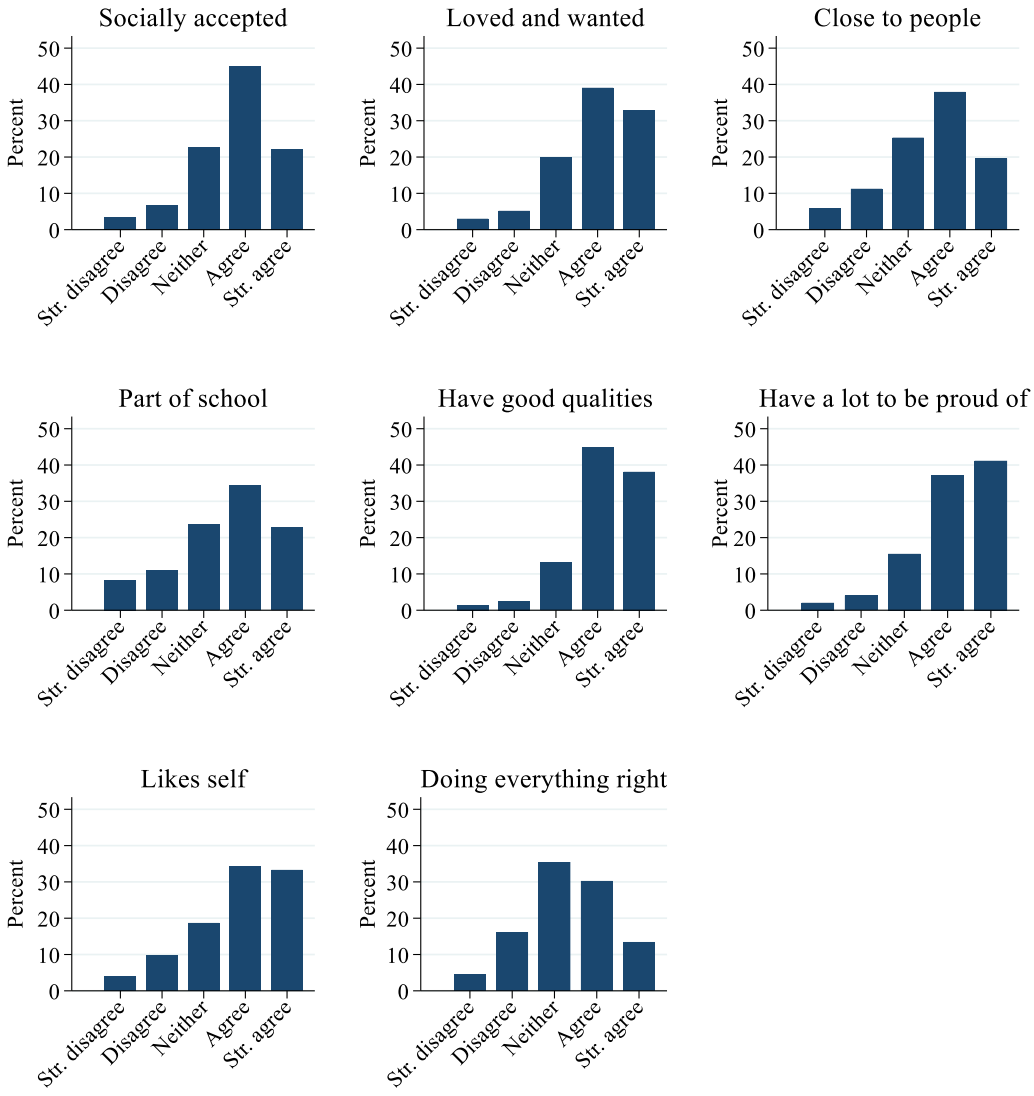
The below figures display the distribution of unweighted responses for the non-cognitive factors I study in this paper. The first figure displays the distribution of the count of "agree" or "strongly agree" responses for the four sense of belonging and four self-worth outcomes. The second figure shows the distribution of responses for the disaggregated non-cognitive survey questions.

Figure A1: Distribution of affirmative responses



Each cell plots the percentage of students who answered in the affirmative ("Agree" or "Strongly Agree") for zero, one, two, three, or all four of their non-cognitive questions. The top row contains the count of affirmative responses for the four sense of belonging questions: I feel socially accepted, I feel loved and wanted, I feel close to people at this school, and I feel like I am a part of this school. The bottom row contains the count of affirmative responses for the four self-worth questions: I have a lot of good qualities, I have a lot to be proud of, I like myself just the way I am, and I feel like I am doing everything just about right. The left column contains the count of affirmative responses for the entire sample of 60,129 students, the middle column contains the count of affirmative responses for the 31,445 female students, and the right column contains the count of affirmative responses for the 28,684 male students. Distributions do not utilize survey weights.

Figure A2: Distribution of individual non-cognitive outcomes



Each cell plots the percentage of students who answered the specified non-cognitive question with each of the five possible answers. Number of observations: 60,129. Distributions do not utilize survey weights.

E. Balance of share female on controls

The below table provides estimates from a regression of the student's share of female peers in their school and grade on the control variables in the baseline model and a vector of school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.

Only one relationship is statistically significant at the 10% level, which is to be expected even if no true correlations exist in the population due to the number of control variables. Furthermore, the size of the correlation is not of practical significance – students above the median age for their grade have a 0.0020 smaller share of female peers.

Table A2: Balance of share female

	Female		Male	
Share female	Coeff.	Std. Err.	Coeff.	Std. Err.
Hispanic or Latino	0.0012	(0.0016)	-0.0010	(0.0013)
Black or African American	-0.0001	(0.0010)	0.0001	(0.0011)
Asian or Pacific Islander	0.0007	(0.0026)	-0.0002	(0.0019)
American Indian or Native American	0.0008	(0.0028)	0.0003	(0.0025)
Other race	0.0011	(0.0028)	0.0004	(0.0021)
Above median age	0.0016	(0.0013)	-0.0020*	(0.0011)
Born in the United States	0.0006	(0.0015)	-0.0005	(0.0015)
Household size				
One	-0.0295	(0.0363)	0.0062	(0.0084)
Two	0.0037	(0.0024)	0.0025	(0.0037)
Four to six	0.0012	(0.0012)	0.0013	(0.0011)
Shelter or group home	0.0051	(0.0064)	0.0016	(0.0040)
Constant	0.5492***	(0.0084)	0.5599***	(0.0092)
Observations	31,445		28,684	

*p < 0.10, ** p < 0.05, *** p < 0.01

F. Baseline estimates with coefficients on control variables shown

Table A3: Baseline regression estimates

	All		Female		Male	
	Belonging	Self-worth	Belonging	Self-worth	Belonging	Self-worth
	(1)	(2)	(3)	(4)	(5)	(6)
Share female	0.295	0.242	-0.00550	-0.197	0.714*	0.633*
	(0.318)	(0.176)	(0.408)	(0.258)	(0.352)	(0.270)
Female	-0.107***	-0.427***				
	(0.0149)	(0.0173)				
Hispanic or Latino	-0.132***	-0.00123	-0.230***	-0.0306	-0.0309	0.0319
	(0.0323)	(0.0333)	(0.0431)	(0.0514)	(0.0413)	(0.0360)
Black or African American	-0.0879**	0.269***	-0.220***	0.336***	0.0655	0.184***
	(0.0334)	(0.0283)	(0.0371)	(0.0336)	(0.0388)	(0.0356)
Asian or Pacific Islander	-0.263***	-0.170**	-0.302***	-0.152*	-0.216***	-0.182**
	(0.0493)	(0.0544)	(0.0769)	(0.0654)	(0.0491)	(0.0566)
American Indian or Native American	-0.379***	-0.167***	-0.350***	-0.126	-0.397***	-0.227***
	(0.0571)	(0.0487)	(0.0705)	(0.0636)	(0.0856)	(0.0621)
Other race	-0.200***	-0.109*	-0.228**	-0.202**	-0.168**	-0.0366
	(0.0451)	(0.0450)	(0.0693)	(0.0765)	(0.0606)	(0.0521)
Above median age	-0.112***	-0.00446	-0.156***	0.00130	-0.0700**	-0.0109
	(0.0172)	(0.0127)	(0.0239)	(0.0203)	(0.0239)	(0.0178)
Born in the United States	-0.0919*	0.0574	-0.0462	0.113**	-0.145**	-0.00229
	(0.0372)	(0.0328)	(0.0511)	(0.0411)	(0.0455)	(0.0394)
Household size						
One	-0.186	-0.224*	-0.301	-0.245	-0.123	-0.223
	(0.111)	(0.0921)	(0.204)	(0.153)	(0.142)	(0.116)
Two	-0.102**	-0.0551	-0.0954	-0.0687	-0.115**	-0.0458
	(0.0333)	(0.0366)	(0.0525)	(0.0483)	(0.0434)	(0.0442)
Four to six	0.0861***	0.0803***	0.0789**	0.0881***	0.0977**	0.0733**
	(0.0201)	(0.0168)	(0.0253)	(0.0248)	(0.0292)	(0.0224)
Shelter or group home	-0.455***	-0.635***	-0.327*	-0.244	-0.550**	-0.869***
	(0.133)	(0.147)	(0.162)	(0.231)	(0.200)	(0.196)
Constant	3.290***	3.115***	3.489***	3.159***	2.883***	2.585***
	(0.160)	(0.104)	(0.211)	(0.145)	(0.189)	(0.169)
Observations	60,129	60,129	31,445	31,445	28,684	28,684

The table contains estimates from regressions of the number of "Agree" or "Strongly Agree" responses to the sense of belonging and self-worth question categories on the share of female peers within a student's school and grade, individual covariates, and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

G. Estimates without controls

Table A4: Regression estimates without controls

	All		Female		Male	
	Belonging	Self-worth	Belonging	Self-worth	Belonging	Self-worth
	(1)	(2)	(3)	(4)	(5)	(6)
Share female	0.335	0.455*	-0.0356	-0.191	0.736*	0.635*
	(0.317)	(0.177)	(0.423)	(0.267)	(0.348)	(0.273)
Constant	3.211***	2.799***	3.503***	3.223***	2.900***	2.626***
	(0.162)	(0.102)	(0.219)	(0.147)	(0.188)	(0.165)
Observations	60,129	60,129	31,445	31,445	28,684	28,684

The table contains estimates from regressions of the number of "Agree" or "Strongly Agree" responses to the sense of belonging and self-worth question categories on the share of female peers within a student's school and grade and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

H. Estimates using average numerical response

Table A5: Regression estimates using average numerical response

	All		Female		Male	
	Belonging	Self-worth	Belonging	Self-worth	Belonging	Self-worth
	(1)	(2)	(3)	(4)	(5)	(6)
Share female	0.215	0.134	-0.0790	-0.204	0.561*	0.486*
	(0.196)	(0.128)	(0.246)	(0.169)	(0.248)	(0.186)
Female	-0.107***	-0.349***				
	(0.0116)	(0.0137)				
Hispanic or Latino	-0.101***	0.00212	-0.160***	-0.00800	-0.0392	0.0149
	(0.0222)	(0.0248)	(0.0294)	(0.0371)	(0.0292)	(0.0274)
Black or African American	-0.0436	0.253***	-0.134***	0.301***	0.0631*	0.191***
	(0.0250)	(0.0232)	(0.0272)	(0.0267)	(0.0305)	(0.0306)
Asian or Pacific Islander	-0.161***	-0.111**	-0.188***	-0.0978*	-0.129***	-0.121**
	(0.0322)	(0.0356)	(0.0457)	(0.0416)	(0.0352)	(0.0406)
American Indian or Native American	-0.248***	-0.110***	-0.229***	-0.101*	-0.267***	-0.134**
	(0.0403)	(0.0310)	(0.0481)	(0.0496)	(0.0623)	(0.0426)
Other race	-0.145***	-0.0377	-0.185**	-0.116	-0.108*	0.0243
	(0.0385)	(0.0379)	(0.0548)	(0.0635)	(0.0495)	(0.0447)
Above median age	-0.0873***	-0.00959	-0.137***	-0.0197	-0.0429*	-0.00267
	(0.0125)	(0.0110)	(0.0175)	(0.0139)	(0.0167)	(0.0153)
Born in the United States	-0.0364	0.0438	0.0102	0.106***	-0.0830*	-0.0200
	(0.0260)	(0.0234)	(0.0350)	(0.0274)	(0.0319)	(0.0311)
Household size						
One	-0.180	-0.183*	-0.296*	-0.157	-0.124	-0.192
	(0.0923)	(0.0838)	(0.141)	(0.106)	(0.123)	(0.117)
Two	-0.0551*	-0.0396	-0.0559	-0.0238	-0.0599	-0.0562
	(0.0253)	(0.0276)	(0.0377)	(0.0375)	(0.0334)	(0.0364)
Four to six	0.0708***	0.0649***	0.0745***	0.0732***	0.0711***	0.0581**
	(0.0140)	(0.0129)	(0.0190)	(0.0179)	(0.0188)	(0.0187)
Shelter or group home	-0.571***	-0.735***	-0.386*	-0.423	-0.693***	-0.922***
	(0.137)	(0.169)	(0.159)	(0.244)	(0.202)	(0.215)
Constant	0.645***	0.463***	0.806***	0.495***	0.324*	0.00495
	(0.0985)	(0.0676)	(0.128)	(0.0919)	(0.135)	(0.108)
Observations	60,129	60,129	31,445	31,445	28,684	28,684

The table contains estimates from regressions of the average numerical response (Strongly Disagree = 1, Disagree = 2, Neither = 3, Agree = 4, Strongly Agree = 5) to the sense of belonging and self-worth question categories on the share of female peers within a student's school and grade, individual covariates, and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

I. Estimates for each survey question

Table A6: Disaggregated regression results for female students

	Female students							
	Socially accepted	Loved and wanted	Close to people	Part of school	Have good qualities	Have a lot to be proud of	Likes self	Doing everything right
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share female	0.0380	-0.129	-0.0914	-0.0816	0.0924	0.0530	-0.461	-0.325
	(0.246)	(0.193)	(0.238)	(0.302)	(0.175)	(0.171)	(0.276)	(0.192)
Hispanic or Latino	-0.0458	-0.111***	-0.161***	-0.216***	-0.0723*	-0.0763*	0.0720	0.0514
	(0.0277)	(0.0286)	(0.0293)	(0.0377)	(0.0284)	(0.0373)	(0.0368)	(0.0364)
Black or African American	0.0723**	0.0465	-0.383***	-0.182***	0.210***	0.202***	0.365***	0.169***
	(0.0217)	(0.0273)	(0.0322)	(0.0362)	(0.0249)	(0.0232)	(0.0310)	(0.0267)
Asian or Pacific Islander	-0.106**	-0.258***	-0.139**	-0.123*	-0.127**	-0.191***	-0.00568	0.0155
	(0.0351)	(0.0407)	(0.0529)	(0.0528)	(0.0391)	(0.0488)	(0.0385)	(0.0356)
American Indian or Native American	-0.151**	-0.204***	-0.173***	-0.235***	-0.0788	-0.0811	-0.0621	-0.0951*
	(0.0470)	(0.0488)	(0.0472)	(0.0509)	(0.0469)	(0.0534)	(0.0463)	(0.0469)
Other race	-0.147*	-0.218***	-0.125*	-0.126*	-0.0196	-0.121*	-0.115	-0.109
	(0.0592)	(0.0572)	(0.0516)	(0.0581)	(0.0555)	(0.0536)	(0.0651)	(0.0669)
Above median age	-0.067***	-0.0596**	-0.149***	-0.182***	-0.059***	-0.074***	0.0332	0.0381*
	(0.0173)	(0.0224)	(0.0217)	(0.0234)	(0.0126)	(0.0149)	(0.0198)	(0.0162)
Born in the United States	0.0388	0.0113	-0.0124	-0.00369	-0.00594	0.0657*	0.163***	0.111***
	(0.0273)	(0.0312)	(0.0451)	(0.0381)	(0.0256)	(0.0264)	(0.0312)	(0.0321)
Household size								
One	-0.180	-0.211	-0.296	-0.300	-0.155	-0.157	-0.121	-0.0600
	(0.142)	(0.138)	(0.205)	(0.206)	(0.127)	(0.0986)	(0.127)	(0.114)
Two	-0.0501	-0.0150	-0.0705	-0.0510	0.0100	-0.0564	0.00663	-0.0351
	(0.0392)	(0.0343)	(0.0422)	(0.0454)	(0.0298)	(0.0320)	(0.0448)	(0.0389)
Four to six	0.0408*	0.0336*	0.0792***	0.0951***	0.0206	0.0431**	0.102***	0.0640**
	(0.0203)	(0.0166)	(0.0200)	(0.0241)	(0.0156)	(0.0157)	(0.0189)	(0.0214)
Shelter or group home	-0.335*	-0.589***	-0.186	-0.177	-0.461*	-0.495*	-0.406	0.0323
	(0.150)	(0.169)	(0.163)	(0.150)	(0.221)	(0.200)	(0.261)	(0.178)
Constant	3.911***	4.751***	4.305***	4.507***	4.283***	4.565***	4.488***	3.627***
	(0.133)	(0.107)	(0.130)	(0.160)	(0.0964)	(0.101)	(0.145)	(0.105)
Observations	31,445	31,445	31,445	31,445	31,445	31,445	31,445	31,445

The table contains estimates from regressions of the probability of affirmative ("Agree" or "Strongly Agree") response to each non-cognitive question on the share of female peers within a student's school and grade, individual covariates, and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Disaggregated regression results for male students

	Male students							
	Socially accepted	Loved and wanted	Close to people	Part of school	Have good qualities	Have a lot to be proud of	Likes self	Doing everything right
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share female	0.341	0.458	0.631*	0.441*	0.121	0.304	0.634**	0.469*
	(0.208)	(0.299)	(0.274)	(0.219)	(0.169)	(0.167)	(0.240)	(0.227)
Hispanic or Latino	0.0150	-0.0365	0.00896	-0.118***	-0.0765*	-0.0354	0.0717**	0.0870**
	(0.0295)	(0.0372)	(0.0301)	(0.0329)	(0.0316)	(0.0284)	(0.0251)	(0.0294)
Black or African American	0.138***	0.0479	0.000402	0.0245	0.167***	0.140***	0.213***	0.0810*
	(0.0310)	(0.0283)	(0.0320)	(0.0379)	(0.0239)	(0.0282)	(0.0322)	(0.0352)
Asian or Pacific Islander	-0.124**	-0.177***	-0.0412	-0.0871*	-0.144***	-0.190***	-0.0419	-0.00585
	(0.0430)	(0.0350)	(0.0394)	(0.0376)	(0.0378)	(0.0365)	(0.0374)	(0.0428)
American Indian or Native American	-0.187***	-0.228***	-0.180**	-0.296***	-0.114**	-0.158***	-0.0622	-0.0875
	(0.0450)	(0.0486)	(0.0587)	(0.0840)	(0.0355)	(0.0441)	(0.0449)	(0.0541)
Other race	0.00611	-0.0810	-0.154*	-0.133	0.0133	0.0159	0.0137	0.0336
	(0.0459)	(0.0506)	(0.0605)	(0.0717)	(0.0434)	(0.0417)	(0.0589)	(0.0498)
Above median age	-0.0160	-0.0382*	-0.0357*	-0.0534*	-0.0202	-0.063***	0.0397*	0.0346*
	(0.0163)	(0.0154)	(0.0160)	(0.0224)	(0.0159)	(0.0167)	(0.0161)	(0.0154)
Born in the United States	-0.0649*	-0.138***	-0.0339	-0.0399	-0.0752**	-0.0372	0.0264	0.0232
	(0.0307)	(0.0339)	(0.0312)	(0.0373)	(0.0283)	(0.0334)	(0.0316)	(0.0323)
Household size								
One	0.0301	-0.210	-0.00477	-0.228	-0.225*	-0.248*	-0.151	0.0200
	(0.108)	(0.120)	(0.125)	(0.145)	(0.0929)	(0.0985)	(0.111)	(0.120)
Two	-0.0300	-0.0443	-0.0876*	-0.0379	-0.0327	-0.0561	-0.0359	-0.0520
	(0.0353)	(0.0352)	(0.0366)	(0.0427)	(0.0299)	(0.0332)	(0.0446)	(0.0400)
Four to six	0.0454*	0.0537**	0.0645**	0.0737**	0.0201	0.0545**	0.0492*	0.0588**
	(0.0188)	(0.0169)	(0.0211)	(0.0234)	(0.0179)	(0.0205)	(0.0207)	(0.0195)
Shelter or group home	-0.615***	-0.831***	-0.392	-0.474*	-0.900***	-0.858***	-0.772***	-0.369*
	(0.180)	(0.200)	(0.203)	(0.182)	(0.200)	(0.192)	(0.180)	(0.175)
Constant	3.928***	4.051***	3.768***	4.115***	4.284***	4.513***	3.712***	2.913***
	(0.117)	(0.168)	(0.162)	(0.114)	(0.0966)	(0.103)	(0.133)	(0.126)
Observations	28,684	28,684	28,684	28,684	28,684	28,684	28,684	28,684

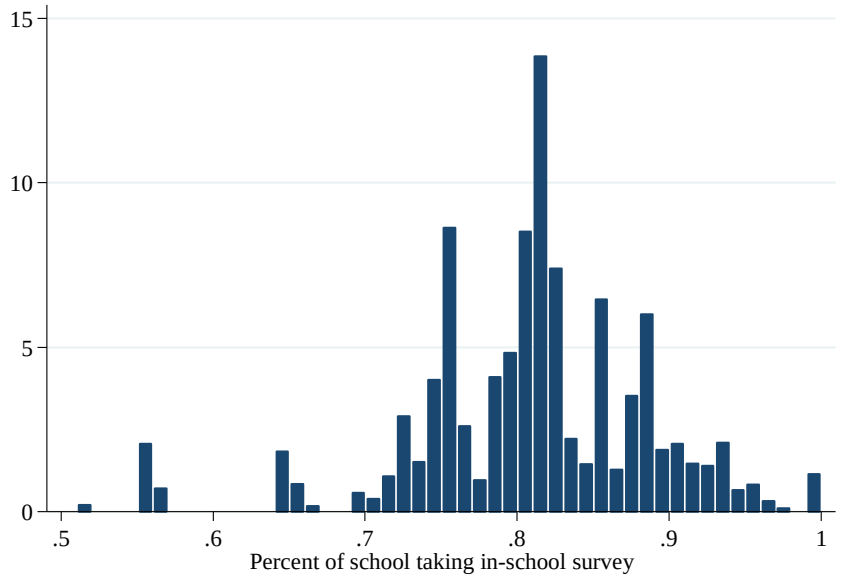
The table contains estimates from regressions of the probability of affirmative ("Agree" or "Strongly Agree") response to each non-cognitive question on the share of female peers within a student's school and grade, individual covariates, and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

J. Sensitivity to measurement error

The primary explanatory variable, the share of female peers, is based on individual-level responses to the in-school questionnaire and is not a precise measure of the true share of female students in a school and grade, as not every student in every school was in attendance or agreed to take the survey. School administrators did report how many students were on their roster for the in-sample grades, allowing me to calculate the percentage of students who responded to the survey. As shown in the histogram below, most students in my sample attended schools that had an 80% response rate or higher, and every student in my sample attended a school with at least a 50% response rate.

Figure A3: Distribution of school response rates



Despite the high responses rates, there is still measurement error in the share of female peers. The effect on the point estimates is unknown ex ante, as the measurement error does not satisfy the classical assumptions. To account for this bias, I follow a multiple imputation procedure and adjust the standard errors according to the formula proposed by Rubin and Schenker (1986).

I assume that the observed share of female students in a school and grade is an unbiased estimate for the true share of female students, which would be true if in-school questionnaire participation is uncorrelated with gender. Next, I estimate the number of missing students in a grade by multiplying the number of missing responses in a school by the share of the grade within each school. I then create 100 simulated data sets filling in the missing students with draws from a Binomial distribution using the probability of being female and number of missing students in each school and grade. I re-estimate my model for each data set, replacing the observed share of female peers with a simulated share of female peers. The average point estimate is a consistent estimate for the theoretical point estimate without measurement error in the share of female

peers, and the average standard error, plus a term to correct for the variation across simulations, is a consistent estimate for the standard error without measurement error.

As shown in the table below, the point estimates for the share of female peers are slightly smaller and the standard errors are slightly larger than in the baseline model. The main results of the paper are unchanged.

Table A8: Regression estimates accounting for measurement error

	All		Female		Male	
	Belonging	Self-worth	Belonging	Self-worth	Belonging	Self-worth
	(1)	(2)	(3)	(4)	(5)	(6)
Share female	0.247	0.197	-0.00112	-0.182	0.613*	0.537*
	(0.315)	(0.192)	(0.398)	(0.268)	(0.367)	(0.276)
Female	-0.107***	-0.427***				
	(0.0149)	(0.0173)				
Control variables	Y	Y	Y	Y	Y	Y
Observations	60,129	60,129	31,445	31,445	28,684	28,684

The table contains estimates from regressions of the number of "Agree" or "Strongly Agree" responses to the sense of belonging and self-worth question categories on the share of female peers within a student's school and grade, individual covariates, and school and grade dummy variables. Standard errors are adjusted for the stratification and clustering of the survey design and for the uncertainty introduced by measurement error in the share of female peers. Observations are weighted by the inverse probability of their selection using weights provided by Add Health.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$