Supporting Information for

Gender brilliance stereotype emerges early and predicts children's motivation in South Korea

Table S1. The questions and scale used in the screener phase in Experiments 1 and 2 (Each option's corresponding scale is displayed in parentheses).

"Smart" screener questions (6 items):

[four relevant questions]

- This child can always answer even the hardest questions from the teacher.
- This child learns things really fast.
- This child can solve very difficult puzzles.
- This child figures things out really quickly.

(after each question) Is this child smart (thumbs up), not smart (thumbs down), or are you not sure (puzzled look)?

[two irrelevant questions]

- This child watches really funny cartoons.
- This child exercises all the time.

(after each question) Is this child smart (thumbs up), not smart (thumbs down), or are you not sure (puzzled look)?

"Nice" screener question:

[four relevant questions]

- This child likes to help other people.
- This child always shares their toys with other children.
- This child tries to make other children feel better when they are sad.
- This child likes to give hugs to family and friends.

Is this child nice (thumbs up), not nice (thumbs down), or are you not sure (puzzled look)?

[two irrelevant questions]

- This child plays on a swing.
- This child likes to listen to music.

Is this child nice (thumbs up), not nice (thumbs down), or are you not sure (puzzled look)?

Scoring: Mean of 6 items (1 = correct answer; 0 = incorrect answer; "smart/nice" to the four relevant questions, and "not sure" and "not smart/not nice" to the two irrelevant questions were coded as correct.)

Scale used in the screener phase:

Table S2. The questions used in the grade task in Experiments 1 and 2. The first two questions were presented with 4 pictures of unfamiliar Asian children (Experiment 1) or White children (Experiment 2) including 2 boys and 2 girls. Participants were then asked the same 2 questions again, except this time they had to choose between 2 verbally presented options ("A boy or a girl?") without viewing pictures.

(1) Who do you think will get the highest score in school/kindergarten?
(2) Who do you think will be the first place in their class?
(3) Who do you think will get the highest score in school/kindergarten? A boy or a girl?
(4) Who do you think will be the first place in their class? A boy or a girl?

An example of the pictures used in the first two questions:

Table S3. Children's own-gender brilliance scores by tasks in Experiment 1 (standard deviations in parentheses).

Age group	Gender	Story task	Guessing task
5	Boys	0.78 (0.31)	0.54 (0.24)
	Girls	0.66 (0.40)	0.60 (0.22)
6	Boys	0.78 (0.36)	0.64 (0.21)
	Girls	0.66 (0.35)	0.56 (0.25)
7	Boys	0.81 (0.25)	0.70 (0.16)
	Girls	0.59 (0.38)	0.54 (0.14)

Note. We submitted children's own-gender brilliance scores obtained from each task (the story task or the guessing task) to a linear regression model with participant gender, participant age, and their interaction as factors. For the story task, the analysis revealed a significant main effect of gender, B = .08, SE = .03, t = 2.24, p = .028. Neither the main effect of participant age, B = -.01, SE = .04, t = -0.18, p = .855, nor the interaction, B = .02, SE = .04, t = 0.55, p = .585, was significant. For the guessing task, the analysis found a significant interaction between participant gender and participant age, B = .05, SE = .03, t = 2.13, p = .035. Neither the main effect of participant gender, B = .03, SE = .02, t = 1.37, p = .174, nor the main effect of participant age, B = .02, SE = .03, t = 0.91, p = .363, was significant.

Table S4. Children's own-gender brilliance scores by tasks in Experiment 2 (standard deviations in parentheses).

Age group	Gender	Story task	Guessing task
5	Boys	0.78 (0.31)	0.71 (0.27)
	Girls	0.72 (0.36)	0.66 (0.25)
6	Boys	0.84 (0.24)	0.69 (0.20)
	Girls	0.72 (0.31)	0.63 (0.19)
7	Boys	0.72 (0.31)	0.68 (0.14)
	Girls	0.44 (0.36)	0.56 (0.19)

Note. We submitted children's own-gender brilliance scores obtained from each task (the story task or the guessing task) to a linear regression model with participant gender, participant age, and their interaction as factors. For the story task, the analysis revealed a significant main effect of participant gender, $B = .08$, $SE = .03$, $t = 2.38$, $p = .020$, and a main effect of participant age, $B = -.09$, $SE = .04$, $t = -2.14$, $p = .035$. However, the interaction was not significant, $B = .05$, $SE = .04$, $t = 1.36$, $p = .178$. For the guessing task, the analysis found a marginally significant main effect of participant gender, $B = .04$, $SE = .02$, $t = 1.79$, $p = .077$. Neither the main effect of participant age, $B = -.03$, $SE = .03$, $t = -1.21$, $p = .230$, nor its interaction with participant gender, $B = .02$, $SE = .03$, $t = 0.60$, $p = .547$, was significant.

Table S5. The four questions used to assess children's interests in Experiment 3 (Each option's numerical score is displayed in parentheses).

(1) Imagine the modi/papu game is right in front of you. Would you want to play the modi/papu game, or would you not want to play it?
[*if "want to play"*] Then, how much do you want to play modi/papu game?
Would you sort of want to play it (= 4), want to play it (= 5), or really want to play it (= 6)?
[*if "not want to play"*] Then, how much do you not want to play modi/papu game?
Would you sort of not want to play it (= 3), not want to play it (= 2), or really not want to play it (= 1)?
(2) Do you like the modi/papu game, or do you not like it?
[*if "like it"*] Then, how much do you like modi/papu game?
Would you sort of like it (= 4), like it (= 5), or really like it (= 6)?
[*if "not like it"*] Then, how much do you not like modi/papu game?
Would you sort of not like it (= 3), not like it (= 2), or really not like it (= 1)?
(3) Imagine you are playing the modi/papu game. Does playing modi/papu game make you happy or sad?
[*if "happy"*] Then, how much playing modi/papu game makes you happy?
Does it make you sort of happy (= 4), happy (= 5), or really happy (= 6)?
[*if "sad"*] Then, how much playing modi/papu game makes you sad?
Does it make you sort of sad (= 3), sad (= 2), or really sad (= 1)?
(4) If you can do something tomorrow, would you play the modi/papu game (= 1) or would you do something else (= 0)?

Figure S1. An example of adult stimuli used in the gender-neutral story task in Experiment 1.

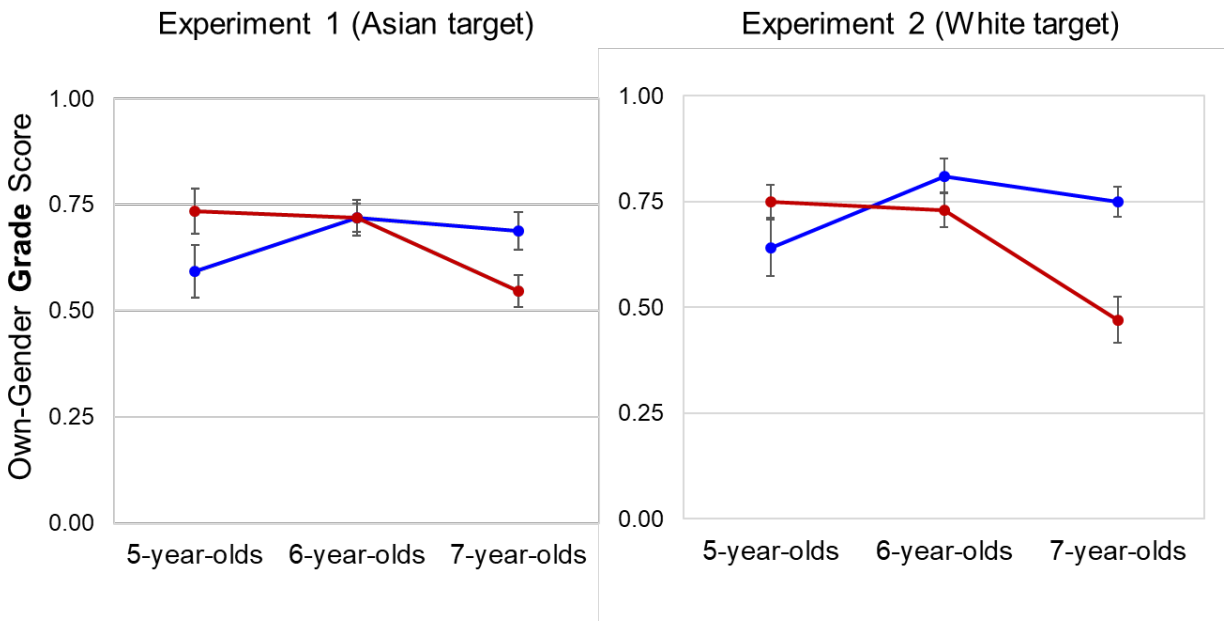

Figure S2. Boys' (blue) and girls' (red) own-gender grade scores in Experiments 1 and 2 by age group (5- vs. 6- vs. 7-year-olds). The error bars represent ± 1 SE.

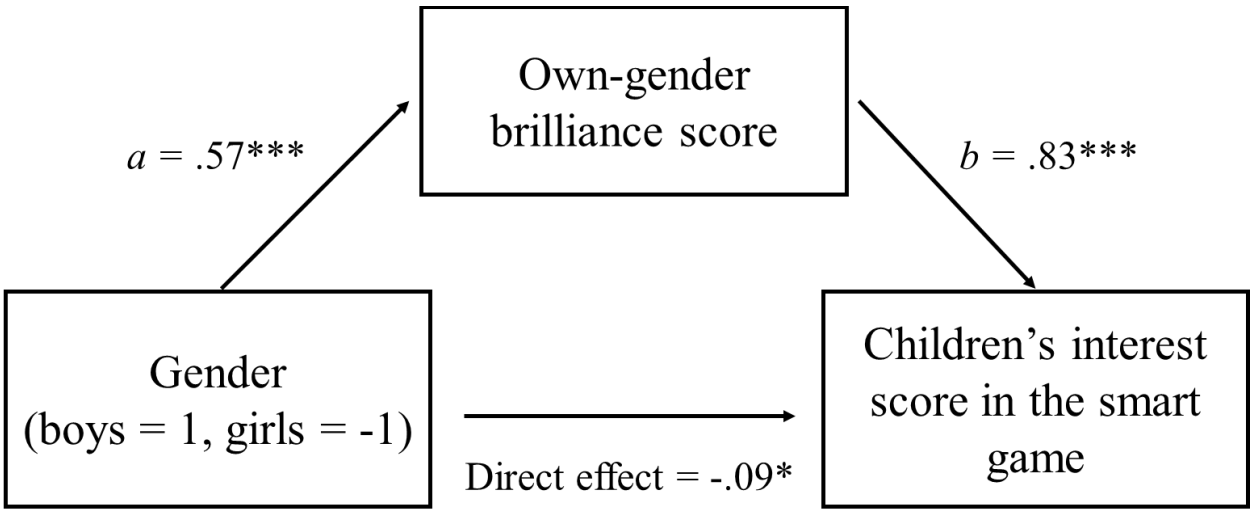

Figure S3. The gender difference in 7-year-olds' interests in the smart game was mediated by their own-gender brilliance scores, indirect effect = .47, 95% CI = [.24, .74]. Unstandardized coefficients are depicted. * $p < .05$, *** $p < .001$

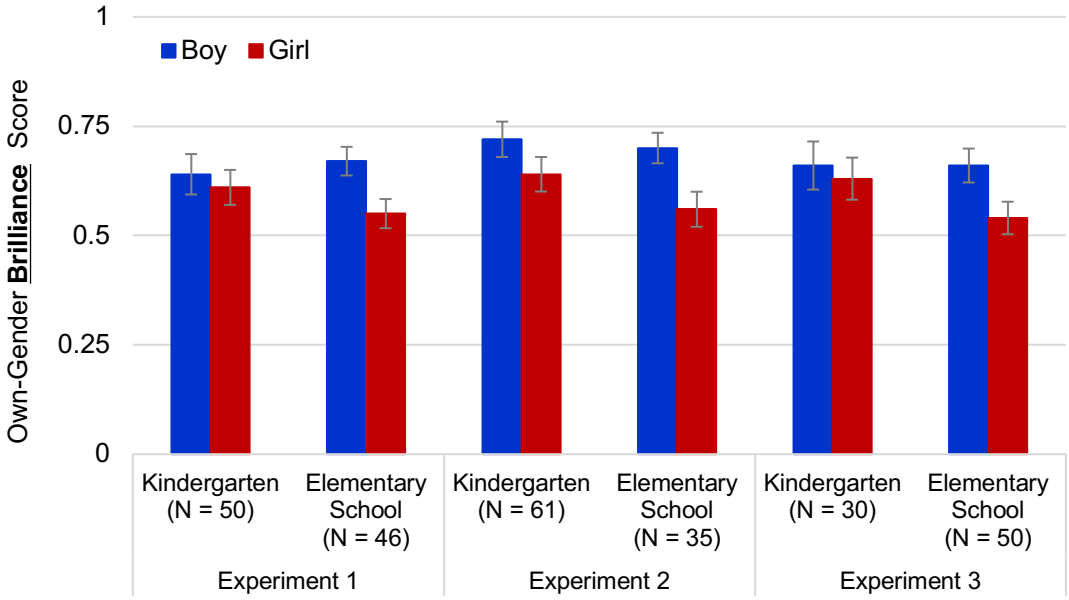

Figure S4. Children's own-gender brilliance scores in all three experiments by institutions they are currently attending (Kindergarten vs. Elementary school). The error bars represent ± 1 *SE*. There was no significant gender difference in kindergartners' own-gender brilliance scores, Exp 1: $F(1, 48) = 0.34$, $p = .565$; Exp 2: $F(1, 59) = 1.82$, $p = .183$; Exp 3: $F(1, 28) = 0.29$, $p = .597$. In contrast, there was a significant gender difference in elementary schoolers' own-gender brilliance scores, Exp 1: $F(1, 44) = 7.58$, $p = .009$; Exp 2: $F(1, 33) = 6.45$, $p = .016$; Exp 3: $F(1, 48) = 5.25$, $p = .026$.

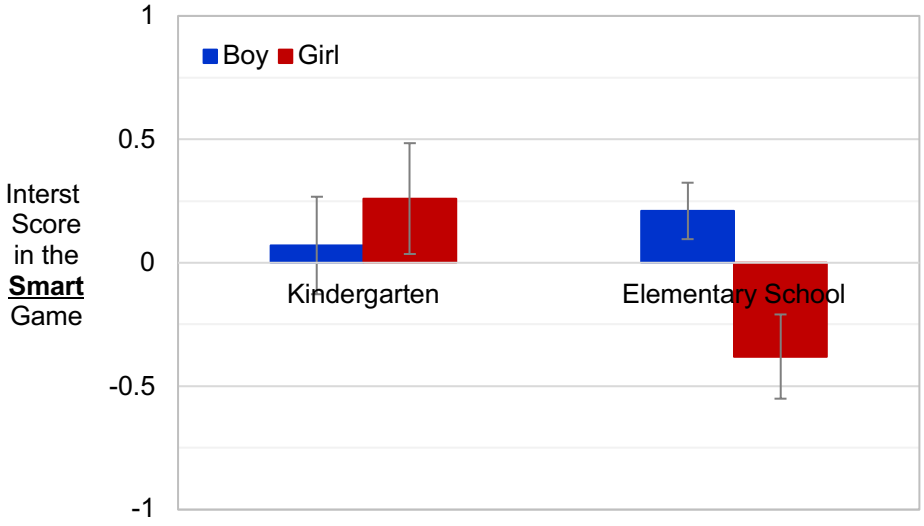

Figure S5. Children's smart game interest scores in Experiment 3 by institutions they are currently attending (Kindergarten vs. Elementary school). The error bars represent ± 1 *SE*. There was a significant gender difference in elementary schoolers, $F(1, 48) = 8.06$, $p = .007$, but not in kindergarteners: $F(1, 28) = 0.39$, $p = .536$.