Falling in Line: Children's Gender Conformity After Feedback Signaling Gender Atypicality

ONLINE SUPPLEMENTAL MATERIAL (OSM)

Table of Contents

Supplementary Appendix A. *Girl and Boy Questions Quizzes (Gender Typicality Threat Manipulation)* 2

Table S1. *Means and SDs for Each Dependent Variable by Age Group, Condition, and Gender* 4

Table S2. *Bivariate Correlations Among All Continuous Measures* 5

Supplementary Appendix B. *Principal Components Analysis of Dependent Measures* 6

Table S3. *Item Loadings from Principal Components Analysis* 8

Supplementary Appendix A

Girl and Boy Questions Quizzes (Gender Typicality Threat Manipulation)

The procedure was age-adapted from Rudman and Fairchild (2004). Children were randomly assigned to Order A or Order B. In **Order A**, they first took the Girl Questions Quiz and learned that they did well on this quiz, then took the Boy Questions Quiz and learned that they did poorly on this quiz. In **Order B**, they first took the Boy Questions Quiz and learned that they did well on this quiz, then took the Girl Questions Quiz and learned that they did poorly on this quiz. As such, girls in Order A received feedback that they were *typical* of their gender (typicality affirmation), girls in Order B received feedback that they were *atypical* of their gender (typicality threat), boys in Order A received feedback that they were *atypical* of their gender (typicality threat), and boys in Order B received feedback that they were *typical* of their gender (typicality affirmation).

The questions that the experimenter read to children are presented verbatim in the screenshots below. Each of the three possible answers was accompanied by a corresponding picture. Following adult typicality threat paradigms (e.g., Vandello & Bosson, 2013), we designed the questions to be difficult enough that children could not easily judge their performance, making our randomly assigned feedback more credible. Because of this, and because the purpose of these quizzes was to manipulate our target variable of interest (i.e., typicality threat vs. affirmation), children's responses were not analyzed. Lastly, question order was fixed to ensure that our manipulation was experienced in exactly the same way by all children in a given condition.

Girl Questions Quiz

Which of these flower fairies is a poppy Fairy?	A B C
Which of these pictures shows a French manicure?	A B C
Which of these wedding dresses has an empire or "Princess" waist?	A B C
Which of these women is wearing a brooch?	A B C
These three women are wearing t-shirts with flowers, but which one is wearing a t-shirt with orchids?	A B C

Boy Questions Quiz

Which of these construction trucks is a bulldozer?	A B C
Which of these football players is Bud Dupree?	A B C
Which of these screwdrivers is a Phillips screwdriver?	A B C
Which of these football teams was the 2016 champion?	A B C
These three men have skateboards, but which one has a penny board?	A B C

Table S1

Means and SDs for Each Dependent Variable by Age Group, Condition, and Gender

			Evaluative Concern		Typicality Enhancement		Atypicality Avoidance	
Age Group	Typicality Condition	Participant Gender	*M*	*SD*	*M*	*SD*	*M*	*SD*
Younger	Affirmation	Girls	−0.35	1.16	0.24	1.04	−0.27	0.73
Younger	Affirmation	Boys	−0.31	1.08	0.18	0.96	0.02	0.96
Younger	Threat	Girls	0.16	0.80	0.68	1.18	−0.27	1.25
Younger	Threat	Boys	−0.19	1.11	0.01	0.80	0.40	1.00
Older	Affirmation	Girls	−0.07	0.92	−0.45	1.01	0.07	0.91
Older	Affirmation	Boys	−0.11	0.98	−0.35	0.88	0.19	0.91
Older	Threat	Girls	0.23	0.83	−0.63	0.73	−0.59	0.75
Older	Threat	Boys	0.58	0.91	0.30	0.81	0.42	1.08

Note. In the analyses reported in the main text, age is treated as a continuous variable, but here we split the sample at the median age (8.16 years) into younger (n = 73) and older (n = 74) groups for the presentation of means and *SD*s by age group, condition, and gender. For all three dependent variables, $M = 0$ and $SD = 1$.

Table S2

Bivariate Correlations Among All Continuous Measures

	1. Age	2. Book of Winners	3. Sticker	4. Fear of Backlash	5. Interest in Same-gender Game	6. Interest in Other-gender Game	7. Top Game Choice	8. Self-stereo-typing	9. Self-esteem
1. Age	——	.25**	−.03	.01	−.13	−.05	−.14	−.30***	−.13
2. Book of Winners	.25**	——	.21*	.14	−.05	−.14	−.04	−.08	−.13
3. Sticker	−.03	.21*	——	.06	−.08	−.13	−.08	.00	−.02
4. Fear of Backlash	.01	.14	.06	——	−.04	.02	−.03	.01	−.35***
5. Interest in Same-gender Game	−.13	−.05	−.08	−.04	——	.24**	.35***	.16	.04
6. Interest in Other-gender Game	−.05	−.14	−.13	.02	.24**	——	.10	−.13	−.05
7. Top Game Choice	−.14	−.04	−.08	−.03	.35***	.10	——	.11	.03
8. Self-stereotyping	−.30***	−.08	.00	.01	.16	−.13	.11	——	.14
9. Self-esteem	−.13	−.13	−.02	−.35***	.04	−.05	.03	.14	——

Note. Bivariate Pearson correlations among all continuous variables, where * $p < .05$, ** $p < .01$, and *** $p < .001$.

Supplementary Appendix B

Principal Components Analysis of Dependent Measures

We submitted our eight measures to a parallel analysis in *R* version 2023.06.1 using the *fa.parallel* function in the *psych* package (Revelle, 2024). This analysis suggested retaining three components. We thus extracted three components in a principal components analysis (PCA) and first applied an oblique rotation (namely, oblimin) to determine the correlations between the components. Given low correlations (*r*s < .09), we ultimately opted for an orthogonal rotation (namely, varimax; Tabachnick & Fidell, 2019).

The component loadings are displayed in Table S3. The items loaded highly onto their respective components (> .50) and did not cross-load onto other components (< .40) (Howard, 2016). Based on the measures that loaded onto each component, we named them *evaluative concern*, *typicality enhancement*, and *atypicality avoidance*.

The *evaluative concern* component comprised two measures: fear of backlash and state self-esteem (which loaded negatively). Past gender development literature has suggested that children as young as 4 years of age demonstrate concerns about appearing gender atypical, which can influence their behavior and self-concepts (e.g., Banerjee & Lintern, 2000; Rudolph & Conley, 2005). Here, the strong negative association between state self-esteem and fear of backlash (see Table S2) suggests that children's concerns regarding what *others* think about their gender (a)typicality is linked to how positively they view themselves (at least in the moment). We thus contend that these two measures combined represent children's concerns about social evaluations when having their gender typicality evaluated or questioned.

The *typicality enhancement* component comprised three measures: participants' interest in the same-gender game, their top choice for a new game, and self-stereotyping. These three

measures were coded such that higher values indicated greater gender conformity. For example, a boy high on this variable would be one who was interested in playing the Boy Questions Game (relative to the Kid Questions Game) again, who selected the Boy Questions Game as his top choice, and reported masculine-stereotyped preferences (e.g., liking action figures but not dolls). As such, we contend that these measures represent a desire for participants to appear more gender-conforming—that is, to enhance their gender typicality.

Lastly, the *atypicality avoidance* component comprised three measures: sticker, book of winners, and participants' interest in the other-gender game (the last of which loaded negatively). Since the Sticker and Book of Winners measures were coded such that higher values = more hiding, this component seems to represent participants' motivation to hide or avoid appearing gender atypical (e.g., hiding one's winner sticker or not wanting to be listed in the book of winners). Moreover, the negative loading of participants' interest in playing the other-gender game indicates that hiding behaviors corresponded with participants' desire *not* to retry a game that might make them appear gender atypical. As such, it appears that—in addition to expressing their gender conformity via typicality enhancement—children may also enact gender conformity by avoiding that which is gender atypical.

We extracted component scores for each participant to use in our main analyses via the default regression-based scoring method. Component scores allowed us to quantify how closely each participant corresponded with the underlying component. After extracting these scores, we checked that their distributions were approximately normal and ensured that there were no missing data.

Table S3

Item Loadings from Principal Components Analysis

	Evaluative Concern	Typicality Enhancement	Atypicality Avoidance
Self-esteem	−.79		
Fear of Backlash	.74		
Interest in Same-gender Game		.78	
Top Game Choice		.73	
Self-stereotyping		.52	
Interest in Other-gender Game			−.69
Sticker			.61
Book of Winners			.54

Note. Items are sorted by factor and loadings less than .40 are suppressed for visual clarity. See the Measures section in the main text for information about how each measure was created.